Oxford Handbooks Online

A Better Self: Freud and Nietzsche on the Nature and Value of Sublimation ⊋

Ken Gemes

The Oxford Handbook of Philosophy and Psychoanalysis
Edited by Richard G.T. Gipps and Michael Lacewing

Abstract and Keywords

The notion of sublimation is essential to Nietzsche and Freud. However, Freud's writings fail to provide a persuasive notion of sublimation. In particular, Freud's writings are confused on the distinction between pathological symptoms and sublimations, and on the relation between sublimation and repression. After rehearsing these problems in some detail it is proposed that a return to Nietzsche allows for a more coherent account of sublimation, its difference from pathological symptoms and its relation to repression. In summary, on Nietzsche's account, while repression and pathological symptoms involve a disintegration (of the self), sublimation involves integration. The chapter provides a brief consideration of some post-Freudian accounts of sublimation that, arguably, represent a return to a more Nietzschean approach. In closing, the chapter contrasts the different bases of Nietzsche's and Freud's valorization of sublimation.

Keywords: sublimation, repression, pathological symptoms, integration, disintegration, master drive, Shreber, Leonardo da Vinci

Introduction

Jacob Golomb writes that '(f)or Freud, sublimation is the only alternative to excessive repression which may deteriorate into an acute neurotic conflict'.[1] He quotes Freud:

> It is precisely in neurotics that … the formation of an ideal heightens the demands of the ego and is met with the most powerful factor favouring repression; sublimation is a way out, by which those demands can be met without involving repression
>
> (Freud 1914: 95).

He then adds: 'This comment accurately reflects the ideas that Nietzsche had formulated on sublimation years before Freud' (Golomb 1989: 69–70). Golomb argues correctly that sublimation is indeed a central notion for both Nietzsche and Freud, but as we shall see, the question of what exactly is involved in sublimation raises profound difficulties. The penultimate sentence in the entry on sublimation in Laplanche's and Pontalis's canonical *The Language of Psychoanalysis* explicitly acknowledges the importance of giving a coherent theory of sublimation:

> In the psychoanalytic literature the concept of sublimation is frequently called upon; the idea answers to a basic need of the Freudian doctrine and it is hard to see how it could be dispensed with
>
> (Laplanche and Pontalis 1973: 433).

One reason that sublimation is a key notion in psychoanalysis is that from Freud's therapeutic point of view successful psychoanalytic treatment aims at sublimation, in as much as sublimation is seen as a necessary condition for psychic health. By bringing to conscious light hitherto repressed drives, desires, and wishes, energy which has previously displayed itself in unpleasurable symptoms may be harnessed and directed to more productive and felicitous ends. And indeed, at first glance, sublimation might seem a clear enough concept. It involves the redirecting of a repressed sexual drive towards a non-sexual aim.[2] As Freud puts it in his paper *On Narcissism*:

> Sublimation is a process that concerns object-libido and consists in the instinct directing itself towards an aim other than, and remote from, that of sexual satisfaction; in this process the accent falls upon deflection from sexuality[3]
>
> (Freud 1914: 94).

Yet Laplanche's and Pontalis's entry on sublimation ends with a starkly negative assessment of attempts to clarify the notion of sublimation: 'The lack of a coherent theory of sublimation remains one of the lacunae in psychoanalytic thought' (Laplanche and Pontalis 1973: 433).

The claim that sublimation is a vexed concept is repeatedly echoed in the 'secondary' literature. Indeed, this is true from the early days of that literature as evidenced in Edward Glover's (1931) assertion in his important paper 'Sublimation, Substitution and Social Anxiety':

> It is generally agreed that prior to 1923 a good deal of confusion existed regarding the exact nature of sublimation. Since then it has increased rather than diminished
>
> (Glover 1931: 263),

More recently Jacques Lacan notes 'the virtually absurd difficulties that authors have encountered every time they have tried to give a meaning to the term "sublimation"' (Lacan 1992: 142–3).

To get an entry to these difficulties let us begin by considering the simple definition given here which emphasizes the redirecting of a drive from a sexual to a non-sexual aim. Before demonstrating the major problem with this definition we shall briefly consider two minor but important problems with this definition. First, there is the problem of distinguishing aims that involve sexual satisfaction and those that do not. After all, it was Freud who argued that many prima facie non-sexual activities (e.g. thumb-sucking) have a sexual component. Perhaps this can be finessed by distinguishing overtly sexual aims from aims that are not overtly sexual. Second, there is the problem that much sublimation seems to involve a diversion of arguably non-sexual instincts. Thus aggressive drives, drives associated with what Freud would later call the death drive, may find sublimated discharge in non-aggressive behaviour.[4]

Substitute Formations, Symptoms, and Sublimations

Now for the major problem: the definition in terms of the substitution of non-sexual for sexual aims fails to distinguish sublimations from symptom formation. An obsessive who compulsively aims to avoid treading on cracks in the pavement may have substituted a non-sexual aim for a sexual aim but that would not normally count as a case of sublimation (though see the following section). On Freud's basic account various drives are repressed and later these may manifest themselves in various behaviours. These behaviours, often labelled by Freud as 'substitute formations', are related in complex ways to the original drive. They are what Freud calls the 'return of the repressed' (Freud 1915: 154). Our key question then is how to separate those substitute formations, those instances of the return of the repressed, which are symptoms, from those that are sublimations?

Of all Freud's writings perhaps the paper that most acutely demonstrates the need for a clear means of separating symptoms from sublimations is his 1910 essay 'Leonardo Da Vinci and a Memory of his Childhood' (Freud 1910a). There Freud argues that Leonardo is a model of that type of individual whose repressed sexual desires find sublimated expression by becoming attached to a powerful drive for scientific and artistic research. In his summarizing final section Freud shows implicit recognition of the difficulties of separating symptoms from sublimations, and, relatedly, illness from health, when he writes:

> We must expressly insist that we have never reckoned Leonardo as a neurotic or a 'nerve case', as the awkward phrase goes. Anyone who protests at our so much as daring to examine him in the light of discoveries gained in the field of pathology is still clinging to prejudices which we have to-day rightly abandoned. We no longer

think that health and illness, normal and neurotic people, are to be sharply distinguished from each other, and that neurotic traits must necessarily be taken as proofs of a general inferiority. To-day we know that neurotic symptoms are structures which are substitutes for certain achievements of repression that we have to carry out in the course of our development from a child to a civilized human being. We know too that we all produce such substitutive structures, and that it is only their number, intensity and distribution which justify us in using the practical concept of illness and in inferring the presence of constitutional inferiority. From the slight indications we have about Leonardo's personality we should be inclined to place him close to the type of neurotic that we describe as 'obsessional'; and we may compare his researches to the 'obsessive brooding' of neurotics, and his inhibitions to what are known as their 'abulias'

(Freud 1910a: 131).

After holding up Leonardo as a model of sublimation throughout the essay, and then in the first sentence of this quotation protesting that he has not reckoned Leonardo to be a neurotic, Freud in the final sentence concludes that we must place Leonardo close to the obsessional neurotic. The distance between sublimation and neurotic symptoms seems vanishingly small. Here we have in microcosm our problem; how to separate sublimations from the neurotic symptoms both of which count as substitutive formations?

The Social Factor in Sublimations

In many of Freud's explications of sublimation there is an emphasis on the social value of the activity resulting from sublimation. Thus in his *New Introductory Lectures* he says:

A certain kind of modification of the aim and a change of object, in which our social valuation is taken into account, is described by us as 'sublimation'

(Freud 1933: 97).

In one of his earliest published references to sublimation in the Dora case of 1905 he talks of the 'the undifferentiated sexual disposition of the child ... being diverted to a higher sexual aim' and thereby providing 'the energy for a greater number of our cultural achievements' (Freud 1905a: 50). In the *Three Essays* dating from the same period he writes that sexual curiosity can be 'diverted (sublimated) in the direction of art' (Freud 1905a:156).

This emphasis on socially valued achievements would provide some means of separating neurotic symptoms from sublimations but at the high cost of introducing a totally non-psychoanalytic, indeed a non-psychological element to the definition; namely that of social valuation. One, presumably uncomfortable, consequence of such an account would be that as social values change, behaviour that was at one time pathological would become a sublimation. For instance, suppose our compulsive crack avoider mentioned earlier successfully presents his activities as a kind of performance art and receives much

social acclaim for his continual performances. Does that move his activity from neurotic symptom to sublimation?

The reference to social acclaim has led many authors to be suspicious of such definitions.[5] In his article 'What is Sublimated in Sublimation' Donald Kaplan begins:

> The whole idea of sublimation has been a vagrant problem for psychoanalysis from the very beginning. This is because sublimation, in one of its meanings, refers to felicitous exercises of the mind, while psychoanalysis is suspicious of any such simple creed
>
> (Kaplan 1993: 549).

One gets a clear picture of the problem posed by the introduction of social valuation as a defining characteristic of sublimation, and the related problem of distinguishing neurotic symptoms from sublimations, by contrasting Freud's essay on Schreber with his essay on Leonardo. The latter is perhaps his most sustained surviving piece covering the topic of sublimation.[6] According to Freud, Schreber's illness stems from his repression of homosexual desires first overtly expressed in his quickly suppressed thought that '(a)fter all it would be nice to be a woman submitting to the act of copulation'. This repression later gave rise to the obsessive thought that God is trying to unman him and turn him into a woman and led to various other obsessive thoughts and behaviour. This would not normally count as a case of sublimation. Contrast this with Freud's Leonardo case. According to Freud, Leonardo, like Schreber, represses his homosexual desires and this repression later leads to his scientific inquisitiveness and inventiveness and his artistic activities including his obsession with capturing in drawings perfect representations of the male body. This for Freud is a classical case of sublimation. Is the only pertinent difference here that in the Schreber case the end activity that results from a repression of erotic desires, namely, obsessive behaviour and thought centring on his relationship to God, is not considered socially valuable, whereas in the Leonardo case the end activity that results from a repression of erotic desires is considered socially valuable? Or, as Kaplan might put it, is it simply that we regard Schreber's exercises of the mind as non-felicitous, but regard Leonardo's as felicitous? Such distinctions as socially valuable/not socially valuable, felicitous/non-felicitous, clearly involve value judgements. Moreover, such distinctions do not seem to be analytical distinctions. Thus it seems a fundamental notion in psychoanalysis is not to be explained in strictly psychoanalytical terms. Furthermore, those terms seem irredeemably normative and this runs against Freud's general claim to be providing a merely scientific/descriptive account of psychological phenomena.[7] Also bringing in the normative results in a questionable form of relativism as norms of what is socially acceptable vary from culture to culture and even between subgroups in a given culture. It is presumably this kind of concern that lead Laplanche and Pontalis to raise the question: 'Should the fact that activities described as sublimated in a given culture are accorded particularly high social esteem be taken as a defining characteristic of sublimation?' (Laplanche and Pontalis 1973: 33).

If we do choose to abandon the social valuation factor in separating sublimation from symptom formation any new account should, I suggest, be tested by how it handles the Leonardo and Schreber cases. The respective behaviours of Leonardo and Schreber may be taken as paradigms of sublimation and pathogenic symptoms. What we are looking for then is a strictly psychoanalytic account which would class Leonardo's behaviour as sublimation and Schreber's as pathogenic symptoms.

The Relation between Sublimation and Repression

Another serious problem raised by Freud's own writings on sublimation concerns the relation between repression and sublimation. In one understanding of Freud's telling of the Leonardo case, Leonardo's homosexual urges are continuously repressed, hence he lives an asexual adult life, and it is this repression that gives rise to his adult scientific and artistic drives:

> the sexual repression which set in after this phase of his childhood (where he experienced the 'excessive tenderness of his mother') caused him to sublimate his libido into the urge to know and established his sexual inactivity for the whole of his later life

> (Freud 1910a: 135).

Here it seems out of place to talk of the repression of the original drives being lifted; for if the repression were lifted there would be no account of the continuation of the scientific and artistic drives and the continual asexuality.[8] On this model, at its simplest, Leonardo's scientific and artistic drives, like his asexuality, are really kinds of symptoms and, as per the usual Freudian model, working through the repressions behind the symptoms should lead to elimination, or possibly the transformation, of the symptoms. The notion that sublimation involves repression is re-enforced by one of Freud's earliest mentions of sublimation in his *Three Essays on Sexuality*.[9] There he refers to 'the diversion of sexual forces from sexual aims and their direction to new ones—a process that deserves the name of "sublimation"' (Freud 1905a: 178). In the same place he goes on to say 'we would place its (sublimation's) beginning in the period of sexual latency of childhood'. Clearly such redirection on the part of the child involves non-conscious repression of those sexual forces qua sexual forces, and such repression is exactly what Freud describes as the progenitor of the period of sexual latency. It is not that the child consciously faces his erotic desires and chooses not to act on them but pursue some other activity as a substitute satisfaction. The sublimation of those desires first requires their repression. Indeed, there is even some tendency in the psychoanalytic literature to at least implicitly identify sublimation as a species of repression. For instance in *Sublimation: Inquiries into Theoretical Psychoanalysis*, Hans Loewald writes 'Traditionally sublimation is classified in psychoanalysis as a defence' (Loewald 1988: 3), and in fact Freud's earliest reference to sublimation in his 1892 letter to Fliess refers to

sublimations as 'protective structures' (Freud 1897: 247). A page later in a footnote Loewald notes that:

> It must be noted, however, that in the psychoanalytic literature, including works by Freud himself, the very distinction between defence and repression has been blurred, in as much as repression is used interchangeably with defense

> (Loewald 1988: 4–5).

So if sublimations are defences and the notions of defences and repressions are interchangeable it follows that sublimations are repressions. Fenichel indeed comments that sometimes sublimation is called a 'successful repression' (Fenichel 1945: 148).

In fact while the Leonardo text and the extract from the *Three Essays on Sexuality* suggest that repression is part of sublimation, often in other texts when Freud explicitly considers repression and sublimation together he claims that sublimation is an alternative to repression. Thus in a passage from *Five Lectures on Psychoanalysis* he says: 'Premature repression makes the sublimation of the repressed instinct impossible; when the repression is lifted, the path to sublimation becomes free once more' (Freud 1910b: 54). In a letter to Putnam, Freud even goes so far as to claim that the two are mutually exclusive: 'It (psychoanalytic theory) teaches that a drive cannot be sublimated as long as it is repressed and that is true for every component of a drive' (Hale 1971: 121).

To clarify the relation between sublimation and repression it helps to consider exactly what gets repressed in repression. Here the best place to turn is Freud's (1915) essay 'Repression' (Freud 1915: 141–58). A crucial point of that essay is Freud's claim that a drive has both an 'ideational' component and an energetic component, what Freud calls 'a quota of affect' (Freud 1915: 152). The ideational component is the intensional content of the drive, including its aim. The energetic component is the force associated with that aim, including the strength of the drive. In repression both the ideational component and the force component are repressed. The repression of the ideational component involves not letting the aim be consciously apprehended. The repression of the force involves not letting the force be expressed in behaviour. While Freud refers to the repression of the ideational component as 'primal repression' he emphasizes that 'the vicissitude of the quota of affect belonging to the representative is far more important than the vicissitude of the idea' (Freud 1915: 153). With this distinction in hand we can now see how sublimation both does and does not involve repression. In sublimation the ideational component may or may not be repressed. In the case of childhood sublimations, the aim of the original drive, the drive's ideational component, is typically held back from conscious apprehension. In the case of sublimation reached through psychoanalytic treatments typically the ideational component becomes available for conscious apprehension. However, in all sublimations, therapeutically achieved or otherwise, the force component is expressed in behaviour. All sublimations involve an expression of a pent-up quota of affect. The picture suggested here is that all sublimations typically take repressions as causal antecedents. In this sense sublimations are another manifestation

Subscriber: Goldsmiths College, University of London; date: 28 February 2019

of the phenomena that Freud calls 'the return of the repressed'. What sublimations undo is the repressing of the energetic component, they steer it to an outlet, an aim that deviates from its original aim. Indeed in the passage quoted earlier from 'Five Lectures on Psychoanalysis' Freud contrasts neurotics who owing to their repressions 'have sacrificed many sources of energy' with cases of sublimation:

> in which the energy of the infantile wishful impulses is not cut off but remains ready for use—the unserviceable aim of the various impulses being replaced by one that is higher, and perhaps no longer sexual
>
> (Freud 1910b: 53–4).

Here Freud is explicitly claiming that in sublimation it is the repression of the energy that is lifted.

This reading does not seem to square with the letter to Putnam where Freud says that sublimation precludes the repression of *any* component of a drive. However, it is worth noting that in that letter he is focused on the clinical practice of psychotherapy. Presumably his idea is that the analyst is faced with a patient presenting various symptoms which the patient experiences as distressing. These pathogenic symptoms are the manifestation of various unconscious repressions. By bringing the repressed drives into full consciousness—this of course involves more than just the patient's intellectual recognition of the drives—this allows for their expression in more acceptable forms. Note, on this model it is not that the ideational component of the drive ceases to be *suppressed* but that the suppression takes on a conscious form that somehow allows the energetic component formerly associated with it to be redirected to new, more acceptable ends. This model does not fit the non-therapeutic cases such as that of Leonardo since there is no conscious recognition on Leonardo's part of his underlying homoerotic drives. Nor does it fit the text of Freud's *Three Essays on Sexuality* where the child in his latency period has no conscious awareness of the desires being sublimated. Clearly the child (unconsciously) represses rather than (consciously) suppresses his erotic desires. So the conclusion to be drawn here is that the sublimation that is the result of successful psychoanalytic intervention involves the lifting of unconscious repression of both the ideational and energetic component; but the sublimation involved in typical cases of scientific and artistic expression or that occurs in the latency period in childhood typically only involve lifting of the repression of the energetic component and continued repression of the ideational component. I suspect it is this model that leads to Fenichel's idea of sublimations as successful repressions; they are successful in that while the ideational component is kept from consciousness the energy associated with it is allowed an outlet. Freud's claim in the Putnam letter that a drive cannot be sublimated as long as any component is repressed would then have to be read as asserting that 'in the practice of psychotherapy a drive cannot be sublimated as long as it is unconsciously repressed'.

Nevertheless, this understanding of the relation of repression and sublimation still leaves us hard-pressed to distinguish sublimation from other instances of the return of the repressed, such as pathological symptoms.

Nietzsche's Solution

Nietzsche, like Freud[10] and like their common predecessor Schopenhauer, takes individual humans to be, at some fundamental level, collections of drives.[11] However, most modern humans, as members of what he denigratingly calls the herd, are simply disorganized collections of competing drives, with different drives having relative ascendancy at different times:

> In the present age human beings have in their bodies the heritage of multiple origins, that is opposite and not merely opposite drives and value standards that fight each other and rarely permit each other any rest. Such human beings of late cultures and refracted lights will on the average be weaker human beings (BGE 200).[12]

In many cases drives, particularly aggressive drives, are treated as per the Freudian model by repression:

> All instincts which are not discharged to the outside are turned back inside. This is what I call the internalization of man. From this first grows in man what people later call his 'soul'. The entire inner world, originally as thin as if stretched between two layers of skin, expanded and extended itself, acquired depth, width, and height to the extent that the discharge of human instinct out into the world was obstructed (GM II 16).

Central to Nietzsche's account is the notion of splitting off. Aggressive drives, which are not viewed as acceptable, typically because acting on them would exact a painful retribution, are suppressed to the point that one does not even acknowledge that one has them; that is to the point where they are not simply suppressed but repressed.[13] They may nevertheless find their outlet often in disguised form. Indeed, often in a form that contradicts their very nature. Thus Nietzsche claims that the Christian value of brotherly love was originally, in fact, a transformed expression of hostile drives to dominate one's fellow men. By successfully preaching brotherly love the weak get their oppressors to voluntarily disarm themselves and become subservient to the values of the weak. In doing so they, both the weak and the strong who have been converted to the values of the weak, split off their contrary aggressive drives from conscious apprehension, so that at the same time they harbour both unacknowledged aggressive drives and acknowledged beneficent drives. This repression and splitting-off of drives makes us sick creatures of what Nietzsche calls *ressentiment*.[14]

For Nietzsche in certain rare cases drives, rather than being split off, are harnessed into a centred, unified, whole. At one point Nietzsche took Wagner to be such a case:

> The dramatic element in Wagner's development is quite unmistakable from the moment when his ruling passion became aware of itself and took his nature in its charge: from that time on there was an end to fumbling, straying, to the proliferation of secondary shoots, and within the most convoluted courses and

often daring trajectories assumed by his artistic plans there rules a single inner law, a will by which they can be explained (UM III 2).

The story of Wagner's achievement of a higher unity borne from some master drive is of course the story Nietzsche would repeat about himself in the dramatic section of *Ecce Homo* where Nietzsche elaborates the subtitle of that work 'How One Becomes What One Is':

> To become what one is, one must not have the slightest notion of what one is ... The whole surface of consciousness—consciousness is a surface—must be kept clear of all great imperatives ... Meanwhile the organizing 'idea' that is destined to rule keeps growing deep down—it begins to command; slowly it leads us back from side roads and wrong roads; it prepares single qualities and fitnesses that will one day prove to be indispensable as a means towards the whole—one by one, it trains all subservient capacities before giving any hint of the dominant task, 'goal', 'aim', or 'meaning' (EH, Why I am So Clever, 9).

This notion of the training of subservient drives is to be explicated in terms of the redirection of those drives away from their initial, primary, goal towards a secondary goal that is more in line with the master drive. This idea is partially expressed in the following passage from *Human, All too Human*:

> *Microcosm and macrocosm of culture*. Man makes the best discoveries about culture within himself when he finds two heterogeneous powers governing there. Given that a man loved the plastic arts or music as much as he was moved by the spirit of science, and that he deemed it impossible to end this contradiction by destroying the one and completely unleashing the other power; then, the only thing remaining to him is to make such a large edifice of culture out of himself that both powers can live there, even if at different ends of it; between them are sheltered conciliatory central powers, with the dominating strength to settle, if need be, any quarrels that break out. Such a cultural edifice in the single individual will have the greatest similarity to the cultural architecture of whole eras and, by analogy, provide continuous instruction about them. For wherever the great architecture of culture developed, it was its task to force opposing forces into harmony through an overwhelming aggregation of the remaining, less incompatible powers, yet without suppressing or shackling them (HAH I, 276).

A point to be emphasized here is that on Nietzsche's ideal weaker drives are not suppressed or shackled. Rather they are to be harnessed to allow their expression in service to a higher aim. Thus in his note books Nietzsche writes:

> Overcoming of the affects? No, if that means their weakening and annihilation. But instead employing them; which may mean a long tyrannizing of them ... At last they are confidently given freedom again: they love us as good servants and

happily go wherever our best interests lie (KSA 12:39, see also HAH I, 224 and 276).

This gives us the material we need to affect a Nietzschean account of the distinction between repression and sublimation. Sublimation is what happens when a drive's primary aim is substituted by a secondary aim that allows for expression of the drive in a manner consonant with the master drive. As John Richardson succinctly puts it: '(D)rive A rules B insofar as it has turned B towards A's own end, so that B now participates in A's distinctive activity' (1996: 33).[15]

Repression is what happens when a drive is denied its immediate aim and is then split off from other drives in the sense that its aims are not integrated with the aims of other drives and it must battle, often unsuccessfully, for any opportunity to achieve expression.

Consider again our cases of Leonardo and Schreber. Leonardo's homosexual sexual drive is redirected towards the secondary aims of scientific, as opposed to sexual, researches and artistic creation, including idealized representations of the male body, as opposed to actual sexual possession of male bodies. These ends fit in with his master drive of scientific and artistic creativity. Schreber, on the other hand, by his own admission, splits off his religious activity completely from his life as senate president. He claims that neither activity in a way informs the other. Schreber does not integrate his sexual drive with his wider life but strives to isolate it and separate its expressions from his other activities.[16]

The Nietzschean solution to the problem of differentiating sublimation from pathological symptoms may be summed up in the slogan *sublimations involve integration or unification, while pathological symptoms involve splittingoff or disintegration*, as we might call it. What is disintegrated is of course the (unified) self. For Nietzsche the difference between repression and sublimation is that in sublimation the stronger drive co-opts a weaker drive as an ally and this allows the weaker drive expression albeit to an end that contains some degree of deflection from its original aim (for example Leonardo's homosexual drive is expressed in possession of idealized representations of male bodies rather than literal possession of male bodies); whereas in repression the stronger drive attempts to stifle any expression of the weaker drive so that its expression is either fully stifled or can only be achieved in a heavily disguised form which often represents the inverse of the original aim (the Christian's hate and envy of, and desire to have power over his fellow man, being expressed as professions of brotherly love and disinterest in power). The relation between sublimation and repression is that often but not always sublimations have repressions as antecedents. If a stronger drive does not at first have sufficient strength to co-opt a weaker drive to it own ends it may simply act to stifle that weaker drive.[17] As the stronger drive gains in strength opportunities for co-option rather than stifling may arise.[18]

Sublimation and *Ressentiment*

For Nietzsche unified selves, what he takes to be genuine persons, are rare achievements; hence he cautions 'one should not at all assume that many humans are "people"'(KSA 12: 491, translation mine), and '(m)ost men present pieces and fragments of man: one has to add them up for a complete man to appear' (KSA 12: 519).[19] It is Nietzsche's aim to foster the development of such genuine persons. In the same vein his Zarathustra says 'it is my art and aim, to compose into one and bring together what is fragment and riddle and dreadful chance' (TSZ II, 21). Sublimation, for Nietzsche, is the primary means to a unified self. Nietzsche, like Freud, takes sublimation as a mark of health.

Nietzsche's important and difficult normative ideas of *amor fati*, eternal recurrence, and affirmation of life are all strongly related to his aim of overcoming *ressentiment*.[20] *Ressentiment* is directly connected to repression in that where there *is ressentiment* there is some drive that we have been forced to stifle. Nietzsche claims that in order to fully love fate or to fully wish back everything eternally, both of which are exemplary ways of affirming life, we would have to overcome all such *ressentiment*. To affirm all of one's life, to overcome *ressentiment*, would be to affirm all of one's drives; life, for Nietzsche, being nothing but a collection of drives.[21] This does not mean to simply let all of one's drives have free expression. That would involve conflict, chaos and, inevitably, disintegration.[22] It means harnessing one's drives to allow them a form of concerted expression, as Nietzsche himself nicely puts the point:

> The multitude and disgregation of impulses and the lack of any systematic order among them results in a 'weak will'; their coordination under a single predominant impulse results in a 'strong' will: in the first case it is the oscillation and lack of gravity; in the later, the precision and clarity of direction (KSA 13: 394).

Sublimation is for Nietzsche the key means to such concerted expression and, hence, to overcoming *ressentiment*.[23]

On Nietzsche's picture of sublimation the questionable element of social valuation drops out of the picture.[24] Another important difference from the Freudian account is that the notion of desexualization also drops out. Where in his account of sublimation Freud deals primarily with a libidinal drive (Eros), Nietzsche allows for a multiplicity of drives. Relatedly, for Nietzsche aggressive drives as well as sexual drives are prone to sublimation. Given the problems of Freud's notion of desexualization, and his concomitant failure to revise his account of sublimation after his positing of a death drive separate from the libidinal drive, these differences in Nietzsche's account are a gain rather than a loss. The price paid for this gain is that Nietzsche, much more than Freud, remains unclear about the range of drives available for sublimation and repression.[25]

On Nietzsche's account the role of the ego in repressions and sublimations is dramatically reduced. Rather than having a higher level agency, such as Freud's ego or superego censor lower level drives, Nietzsche simply posits an agonal struggle between the drives; thus he says in *Beyond Good and Evil*: 'The will to overcome an affect is, in the end, itself only the will of another, or several other, affects' (BGE 117). And even in those cases where there is consciousness of the drive that is causing a disturbance, that consciousness is not the engine of repression:

> (*T*)*hat* one *wants* to combat the vehemence of a drive at all, however, does not stand within our own power; nor does the choice of any particular method; nor does the success or failure of this method. What is clearly the case is that in this entire procedure our intellect is only the blind instrument of *another drive*, which is a *rival* of the drive whose vehemence is tormenting us … While 'we' believe we are complaining about the vehemence of a drive, at bottom it is one drive *which is complaining about the other*; that is to say: for us to become aware that we are suffering from the *vehemence* of a drive presupposes the existence of another equally vehement or even more vehement drive, and that a *struggle* is in prospect in which our intellect is going to have to take sides (D: 109).

While this passage allows that our conscious selves may be aware of the drive that is causing us distress, it emphasizes another drive, not the conscious I, as the repressing force. Other texts from Nietzsche make clear that he envisages such struggles between drives often occurring with no conscious awareness, that is with no participation of the conscious I, or ego, as Freud calls it (cf. for instance, the passages quoted earlier describing Wagner and Nietzsche himself).

Klein, Segal, and Loewald: Towards a Reconciliation of Freud and Nietzsche

While most of Freud's texts on sublimation emphasize notions such as desexualization and social valuations, there are occasional texts that emphasize the unifying element in sublimations. In *The Ego and the Id* Freud tells us that sublimated energy may 'retain the main purpose of Eros—that of uniting and binding—as it helps towards establishing unity, or a tendency to unity' (Freud 1923: 45).[26] While Freud fails to thematize this notion of sublimation as unification, some post-Freudians, in particular those working in the tradition of Melanie Klein, do attempt to thematize the important notions of disintegration and integration in their discussions of sublimation.

In her essay 'Infant Analysis' Klein explicitly turns to consideration of Freud's analysis of Leonardo in order to 'set forth more exactly the analogies and differences between symptoms and sublimations' (Klein 1926: 41). The key component she attributes to Leonardo which allows him to achieve sublimations rather than hysterical conversions is his capacity for 'identification with the objects of the world around him'. It is this which allows his fixations to 'become consonant with the ego' (Klein 1926: 43). According to

Klein, having introjected various objects which are assigned both good and bad qualities—this for Klein is a feature of what she calls the depressive position—sublimation is achieved when hostility to introjected objects is overcome through phantasies that allow a symbolic representation of the introjected object in a non-hostile form. It is through such phantasies and symbolic representations that these introjected objects are integrated with the ego. While we cannot hope to fully rehearse here Klein's notion of depressive position etc. what is clear is that for Klein sublimation involves integration and overcoming of hostility to inner objects. Since these inner objects are for Klein manifestations of drives (for instance bad objects are manifestations of aggressive drives), Klein's account of sublimation bears a striking resemblance to the Nietzschean account offered earlier.[27]

The Kleinian account of sublimation as integration is a theme taken up by Hanna Segal in her 'A Psycho-Analytical Approach to Aesthetics' which concludes that 'a satisfactory work of art is achieved by a realization and sublimation of the depressive position' (Segal 1952: 206). For Segal this sublimation involves 'the desire to unite' (Segal 1952: 207). Art, for Segal is a form of sublimatory activity, a working through of the depressive position, that allows us through symbolic representations to take objects towards which we have sadistic hostile impulses and reintegrate them into a world which is 'whole, complete and unified' (Segal 1952: 204).

Another psychoanalyst, influenced by Klein, who identifies sublimation with a type of integration, albeit under the name of 'internalization', is Hans Loewald. Internalization according to Loewald involves 'unconscious ego processes that undo the splitting off' that is characteristic of repression (Loewald 1973: 12). The splitting off that characterizes repression, according to Loewald, 'maintains psychic processes and structures on lower organizational levels' while sublimation, or internalization as he calls it, 'leads to higher organization and an enriched psychic life' (1973: 14). In his later book *Sublimation*, in reference to Freud's Leonardo case, Loewald notes that 'there are unitary experiences which give way to experiences of differentiation, but that in sublimation the experience of unity is restored' (Loewald 1988: 45). Interestingly, Loewald (1973) cites Nietzsche, and in particular Nietzsche's *Genealogy of Morals*, as a precursor for his own notion of internalization. In fact Nietzsche's term 'internalization' is closer to the notion of repression than sublimation. For Nietzsche, as expressed in the passage from GM II 16 quoted earlier, internalization is the turning back of instincts that cannot be discharged externally.

What is missing from these post-Freudian accounts of sublimation as integration or unification, or a higher level of organization, is an account of what exactly is meant by integration, unification, or higher organization. Nietzsche's account of a master drive with a determinate aim realigning the aims of weaker drives, to aims which augment rather than conflict with the aim of the master drive, at least provides a start to such an account. Whether Kleinians or other post-Freudians can utilize such a Nietzschean conception of sublimation as unification without joining Nietzsche in putting such strong conditions on what is necessary for unification, or without joining Nietzsche in his

rejection of the notion of the role of the ego, is something that awaits further investigation.

The Value of Sublimation

Above it was mentioned that both Nietzsche and Freud value health and took sublimation to be a means toward health. Yet where Freud's notion of health involves a fairly moderate normative agenda, Nietzsche's notion of health involves a much more radical normative agenda. Freud, working more within a traditional medical model, took his therapy to be primarily aimed at relieving psychic distress. In this vein he spoke of 'transforming … hysterical misery into common unhappiness' (Freud 1893–5: 305). For Freud, repression leads to anxiety—in his quasi-energetic model this is expressed in terms of the affective state caused by inhibition of discharge; pent up energy denied realization causes a kind of system instability. Here it is clear why, given the aim of a reduction of misery, sublimation, which allows for discharge of the drives, is to be preferred to repression. Continuing in terms of the energetic model, repression leads to a building up of pressure, sublimation allows for a release of pressure.

However on Nietzsche's evaluative scheme the preference for sublimation over repression is less clearly grounded. Nietzsche certainly did not place a high value on contentment and often inveighed against it as a form of complacency; what he disparagingly refers to as 'green pasture happiness of the herd' (BGE 44)—implying that such a goal is worthy only of bovine animals. Where Nietzsche places the pursuit of happiness as the aim of the last men and other lower types (Z Prologue 5), Freud, to some extent echoing Schopenhauer, postulates it as the aim of humans in general:

> They strive for happiness; they want to become happy and to remain so. This endeavour has two sides, a positive and a negative aim. It aims, on the one hand, at an absence of pain and unpleasure, and, on the other, at the experiencing of strong feelings of pleasure
>
> (*Civilization and Its Discontent*, Freud 1930: 76).

Against this type of characterization Nietzsche, presumably having in mind various utilitarian philosophers, quipped: 'Man does not strive after happiness; only the Englishman does that' (TI, Maxims and Arrows, 12).

In *Zarathustra* Nietzsche refers contemptuously to 'the last man' as the 'most contemptible man' who has 'discovered happiness' and who 'makes everything small'; such a man 'will no more shoot the arrow of his longing out over mankind and the string of his bow will have forgotten how to twang' (Z, Prologue, 5). This metaphor of the dissipation of tension contrasts vividly with his characterization of highly repressed individuals who seethe with ressentiment. Indeed one of the most startling of his characterizations of the generally negatively viewed psychic phenomena of ressentiment is that it helped make man an 'interesting animal', that it gave him depth and that

without it '(h)uman history would be much too stupid an affair' (GM I 6–7). There is a touch of both admiration and horror when Nietzsche writes of 'the ressentiment of beings denied the true reason, that of deed', and of occasions where 'ressentiment itself becomes creative and gives birth to values' (GM I 10).

One might think that perhaps Nietzsche's objection to ressentiment and repression was more on a quasi-moral ground: that one overtaken by ressentiment and repression typically has a false view both of the object of ressentiment and of one's own nature. For instance, Nietzsche claims that repressed resentful slaves imagine, incorrectly, that their masters are free to act other than in a masterly (i.e., slave oppressive) way and that they, the slaves, have freely chosen to be meek, mild, turn the other cheek, etc., 'weakness is to be lied into a merit' (GM I 14). Now for those who, working under the influence of Christian morality, place a categorical value on truth, the objection that repression and ressentiment typically lead their bearers to a false view of themselves and others, counts as probative. In the *Genealogy of Morality* Nietzsche paints a story of repressed slaves becoming creatures of resentment involved in massive self-deception knowing that his readers, as inheritors of Christian morality, will naturally take the charge of self-deception as a severe criticism. But what about those, allegedly such as Nietzsche himself, who strive for an *arche* beyond Christian morality? The demand to be free of self-deception, the demand for truth at any price, does not sit well with Nietzsche's repeated claims that lies and myths are a necessary part of life. Nietzsche repeatedly writes of the need for self-deception (BGE 2, 4) and even claims that ignorance of one's deeper drives and motivations can often be healthy phenomenon (EH, Why I am So Clever 9). Christian morality treats truth as a categorical and ultimate value, that is, as a final value that is higher than all other values. For that reason it espouses a total rejection of all that is opposed to truth, including falsity, myth, and self-deception. While Nietzsche does indeed value truth, he sees it as one of many competing values. Health for Nietzsche involves a certain balancing of these values; taking truth as a value to which everything must be subordinated is for Nietzsche a form of pathology. Nietzsche recognizes that the health of a certain individual or indeed of a certain culture is sometimes built upon certain lies, myths, and self-deception. A Nietzsche who valorized truth above all else would be subject to Nietzsche's own trenchant critique of the ascetic ideal and its fetishization of truth (GM III, *passim*).

It seems reasonable to conclude that Nietzsche does not prefer sublimation over repression because the former facilitates a reduction of misery, and an easing of psychic tension, or because repression leads to false beliefs or even self-deception.

One strand of Nietzsche's arguments against repression is that it prevents a certain kind of unity that he seems to value.[28] It is in this deeper sense that through the effects of repression we become 'strangers to ourselves' (GM, Prologue, 1); we do not simply lack knowledge of ourselves but more importantly we have become estranged from parts of ourselves. Through repression parts of our selves are split-off, not simply in the sense that we have no conscious access to them, but in the sense that we contain within us hidden affects and drives. These are separate movers that are not part of any integrated

whole. Taken to the extreme, this notion of being strangers to ourselves actually threatens the notion of a unified self. That is to say, we have strangers *within* ourselves, so that, in fact, our self is no genuine self. We are nothing more than a jumble of different voices/drives having no overall unity.[29] But what is the basis for valorizing unity? Is it merely an inheritance from the German romantics? Gemes (2006a) argues that Nietzsche values unity because he thinks that it is a necessary condition for freedom and agency. But then why value freedom and agency? Are these some of Nietzsche's ultimate, categorical values?[30]

Alternatively, or additionally, one might argue that Nietzsche, following the romantics, believed that unity is the precondition for great creativity and high culture; these being (some of) Nietzsche's highest values. But Nietzsche himself, along with his story of the repressed, resentful slaves/priests, and his story of Socrates, who create whole new moral universes, seems to run counter to this claim. One suspects that Harold Bloom's thesis that strong poets, and, more generally, other creative types, have a sort of resentment of their predecessors, a certain 'anxiety of influence', which spurs them to their greatest creative deeds, would not at all be uncongenial to Nietzsche. One might argue that in many such cases great creativity comes with the price of a certain denigration of life—for Nietzsche the Christian and Socratic denigration of the non-rational drives is simply a denigration of life. So even allowing that both repression and ressentiment can be spurs to great creativity, Nietzsche would prefer a form of creativity rooted in sublimation as it is more easily wedded to a genuinely life-affirming spirit.

Here is my final suggestion: One of Nietzsche's fundamental grounds for preferring sublimation to repression is ultimately aesthetic. According to Nietzsche repressed persons have parts of themselves that they cannot accept; thus fundamentally ressentiment is really directed at those very parts of themselves. What the repressed person of ressentiment really hates is those parts of himself that his weakness, his inability to act, has led him to repudiate and deny. This self hatred is ultimately disfiguring and ugly. It is intriguing that Nietzsche suggests that his fundamental objection to Socrates was that he was ugly. For Nietzsche, Socrates is a monster of reason who has disfigured himself literally and metaphorically by suppressing all that is not reason in himself:

> In origin, Socrates belonged to the lowest class: Socrates was plebs. We know, we can still see for ourselves, how ugly he was. But ugliness, in itself an objection, is among the Greeks almost a refutation. Was Socrates a Greek at all? Ugliness is often enough the expression of a development that has been crossed, *thwarted* by crossing. Or it appears as declining development. The anthropologists among the criminologists tell us that the typical criminal is ugly: *monstrum in fronte, monstrum in animo* (A, 'The Problem of Socrates', 3).[31]

In the *Gay Science* (290) Nietzsche extols a certain kind of self-satisfaction as the 'one needful thing'. In doing so he raises a distinctly aesthetic objection to that human being

who 'is dissatisfied with himself and is constantly ready for revenge', namely, that we, his potential victims, have to 'endure his ugly sight'.

It is doubtful that Nietzsche had some single ultimate categorical (i.e. end in itself) value. More likely he had a variety of categorical values. Among other things, he valued, as ends in themselves, great individuals, high culture, affirmation of life, expression of power, and various aesthetic values (for instance, beauty). Indeed many of these he valued both categorically and hypothetically (for the things they lead to). For instance, great individuals are valued both as ends in themselves and as a means towards the development of high culture. His strong preference for sublimation over repression is perhaps best understood as embodying his belief that the former provides a more likely path than the latter to realizing these values.[32]

References

Adler, A. (1928). *Über den nervösen Charakter: Grundzüge einer vergleichenden Individual-Psychologie und Psychotherapie*. Munich: Verlag von J. F. Bergmann.

Adler, N. (1986). 'Sublimation and Addiction', *Psychoanalytic Psychology 3*: 187–191.

Arlow, J. (1955). 'Sublimation', *Journal of the American Psychoanalytic Association 3*: 515–527.

Assoun, P. (2000). *Freud and Nietzsche*. London: Continuum.

Bergler, E. (1945). 'On a Five-Layer Structure in Sublimation', *The Psychoanalytic Quarterly 1*: 76–97.

Bos, J. (1992). 'On the Origin of the Id', *International Review of Psycho-Analysis 19*: 433–443.

Chapman, A. and Chapman-Satana, M. (1995). 'The Influence of Nietzsche on Freud's Ideas', *British Journal of Psychiatry 166*: 251–253.

Deri, F. (1939). 'On Sublimation', *The Psychoanalytic Quarterly*, *8*: 325–334

Fenichel, O. (1945). *The Psychoanalytic Theory of Neurosis*. New York: Norton.

Freud, S. (1893–5). *Studies on Hysteria*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 2. London: Hogarth Press.

Freud, S. (1897). 'Letter 61'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 1 (pp. 247–248). London: Hogarth Press.

Freud, S. (1905a). 'Three Essays on Sexuality'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 7 (pp. 123–243). London: Hogarth Press.

Freud, S. (1905b). 'Fragment of an Analysis of a Case of Hysteria' In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 7 (pp. 1–122). London: Hogarth Press.

Freud, S. (1910a). 'Leonardo Da Vinci and a Memory of his Childhood'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 11 (pp. 57–137). London: Hogarth Press.

Freud, S. (1910b). 'Five Lectures on Psychoanalysis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 11 (pp. 3–55). London: Hogarth Press.

Freud, S. (1914). 'On Narcissism'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 67–102). London: Hogarth Press.

Freud, S. (1915). 'Repression'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 141–158). London: Hogarth Press.

Freud, S. (1920). 'Beyond the Pleasure Principle'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 18 (pp. 1–64). London: Hogarth Press.

Freud, S. (1923). 'The Ego and the Id'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 19 (pp. 1–66). London: Hogarth Press.

Freud, S. (1930). 'Civilization and its discontents'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 21 (pp. 57–146). London: Hogarth Press.

Freud, S. (1933). '*New Introductory Lectures on Psychoanalysis*'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 22 (pp. 1–193). London: Hogarth Press.

Gardner, S. (1993). *Irrationality and the Philosophy of the Psychoanalysis*, Cambridge: Cambridge University Press.

Gemes, K. (2006a). 'Nietzsche on the Sovereign Individual, Free Will and Autonomy', *Aristotelian Society* Suppl. *80*: 321–339.

Gemes, K. (2006b). '"We Remain of Necessity Stranger to Ourselves": The Key Message of Nietzsche's Genealogy'. In C. D. Acampora (ed.), *Nietzsche's* On the Genealogy of Morals: *Critical Essays* (pp. 191–208). Lanham, MD: Rowman & Littlefield.

Gemes, K. (2008). 'Nihilism and the Affirmation of Life: A Review of and Dialog with Bernard Reginster', *European Journal of Philosophy 16*: 459–466.

Gemes, K. (2009). 'Freud and Nietzsche on Sublimation', *Journal of Nietzsche Studies 38*: 38–59.

Glover, E. (1931). 'Sublimation, Substitution and Social Anxiety'. *International Journal of Psycho-Analysis 12*: 263–297.

Golomb, J. (1989). *Nietzsche's Enticing Psychology of Power*. Jerusalem: The Magnes Press.

Golomb, J., Santaniello, W., and Lehrer, R. (eds) (1999). *Nietzsche and Depth Psychology*. Albany, NY: State University of New York Press.

Hacker, F. (1972). 'Sublimation Revisited', *International Journal of Psycho-Analysis 53*: 219–223.

Hale, N. (ed.) (1971). *James Jackson Putnam and Psychoanalysis*. Cambridge, MA: Harvard University Press.

Hartmann, H. (1948). 'Comments on the Psychoanalytic Theory of Instinctual Drives', *The Psychoanalytic Quarterly 17*: 368–388.

Hartmann, H. (1955). 'Notes on the Theory of Sublimation', *Psychoanalytic Study of the Child 10*: 9–29.

Heidegger, M. (1979). *Nietzsche*, Vol. I: *The Will to Power as Art*, trans. D. F. Krell. New York: Harper and Row.

Huddleston, A. (2014). 'Nietzsche's Meta-axiology: Against the Skeptical Readings', *British Journal for the History of Philosophy 22*: 322–343.

Kaplan, D. (1993). 'What is Sublimated in Sublimation', *Journal of the American Psychoanalytic Association 41*: 549–570.

Kaufmann, W. (1980). *Nietzsche: Philosopher, Psychologist, Antichrist*, 4th ed. Princeton: Princeton University Press.

Kaywin, L. (1977). 'Problems of Sublimation', *Journal of the American Psychoanalytic Association 25*: 363–370.

Klein, M. (1923). 'The Development of a Child', *International Journal of Psycho-Analysis 4*: 419–474.

Klein, M. (1926). 'Infant Analysis', *International Journal of Psycho-Analysis 7*: 31–63.

Klein, M. (1930). 'The Importance of Symbol-Formation in the Development of the Ego', *International Journal of Psycho-Analysis 11*: 24–39.

Kris, E. (1955). 'Neutralization and Sublimation—Observations on Young Children', *Psychoanalytic Study of the Child 10*: 30–46.

Lacan, J. (1992). *The Seminar of Jacques Lacan Book VII: The Ethics of Psychoanalysis 1959–60*, ed. J. Miller. New York: Norton.

Laplanche, J. and Pontalis, J. (1973). *The Language of Psychoanalysis*. New York: Norton.

Lehrer, R. (1995). *Nietzsche's Presence in Freud's Life and Thought: On the Origins of a Psychology of Dynamic Unconscious Mental Functioning*. Albany, NY: State University of New York Press.

Loewald, H. (1962). 'Internalisation, Separation, Mourning, and the Superego', *The Psychoanalytic Quarterly 31*: 483–504.

Loewald, H. (1973). 'On Internalization', *International Journal of Psycho-Analysis 54*: 9–17.

Loewald, H. (1988). *Sublimation*: *Inquiries into Theoretical Psychoanalysis*. New Haven, CT: Yale University Press.

May, S. (1999). *Nietzsche's Ethics and his War on 'Morality'*. Oxford: Clarendon Press.

Muller, J. (1999). 'Modes and Functions of Sublimation', *Annual of Psychoanalysis 26*: 103–125.

Nietzsche, F. (1966). (BGE) *Beyond Good and Evil*, trans. W. Kaufmann. New York: Vintage.

Nietzsche, F. (1969). (EC) *Ecce Homo*. In *On the Genealogy of Morals and Ecce Homo*, trans W. Kaufmann and R. J. Hollingdale. New York: Vintage.

Nietzsche, F. (1969). (GM) *On the Genealogy of Morals*. In *On the Genealogy of Morals and Ecce Homo*, trans. W. Kaufmann and R. J. Hollingdale. New York: Vintage.

Nietzsche, F. (1975). (TSZ) *Thus Spoke Zarathustra*, trans. R. J. Hollingdale. Middlesex: Penguin.

Nietzsche, F. (1983). (UM) *Untimely Meditations*, trans. R. J. Hollingdale. Cambridge: Cambridge University Press.

Nietzsche, F. (1986). (UM) *Human, All Too Human*, trans. R. J. Hollingdale. Cambridge: Cambridge University Press.

Nietzsche, F. (1988). (KSA) *Sämtliche Werke: Kritische Studienausgabe in 15 Einzeländen*, eds G. Colli and M. Montinari. Berlin: de Gruyter.

Nietzsche, F. (1997). (D) *Daybreak*, eds M. Clark and B. Leiter, trans. R. J. Hollingdale. Cambridge: Cambridge University Press.

Reginster, B. (1997). 'Nietzsche on Ressentiment and Valuation', *Philosophy and Phenomenological Research 57*: 281–305.

Richardson, J. (1996). *Nietzsche's System*. Oxford: Oxford University Press.

Roheim, G. (1943). 'Sublimation', *The Psychoanalytic Quarterly 12*: 338–352.

Sandler, J. and Joffe, G. (1966). 'On Skill and Sublimation', *Journal of the American Psychoanalytic Association 14*: 335–355.

Sartre, J. (1958). *Being and Nothingness*. London: Hazel Barnes.

Schacht, R. (1983). *Nietzsche*. London: Routledge & Kegan Paul.

Segal, H. (1952). 'A Psychoanalytical Approach to Aesthetics', *International Journal of Psycho-Analysis 33*: 196–207.

Segal, H. (1957). 'Notes on Symbol Formation', *International Journal of Psycho-Analysis 38*: 391–397.

Segal, H. (1991). *Dream, Phantasy and Art*. London: Routledge.

Sharpe, E. (1930). 'Certain Aspects of Sublimation and Delusion', *International Journal of Psycho-Analysis 11*: 12–23.

Volney, G. (1992). *Freud on Sublimation—Reconsiderations*. Albany, NY: State University of New York Press.

Notes:

(1) This chapter is a development of Gemes (2009).

(2) Freud often distinguishes between the aim and object of a drive. On one version of Freud's aim/object contrast the aim of a drive is simply satisfaction, or put another way, its aim is to discharge its associated quanta of energy; and the object of a drive is that characteristic activity or things towards which that energy is directed. However, Freud often uses the aim/object contrast to differentiate the characteristic activity of a drive (the aim) from the particular things that drive focuses on at different times (the objects). In this chapter the term 'aim' is typically used in the sense of the active direction of a drive. Thanks are due to Sebastian Gardner for bringing my attention to this point.

(3.) The term Freud invariably used is '*Trieb*' which the translators of the Standard Edition have unfortunately rendered as 'instinct' rather than 'drive'.

(4.) Bergler (1945: 77) similarly notes:

At the time Freud formulated his views on the subject (of sublimation), only the repressed phallic and pre-genital wishes and their contributories were considered to be part of the id. Later, Freud put at least equal stress on repressed aggressive trends. However, no revision of the problem of sublimation was undertaken on the basis of this inclusion of aggressive trends.

(5.) Cf., for instance, Glover (1931: 266–7); Hartmann (1955: 10); Arlow (1955: 515).

(6.) The editors of the Standard Edition hypothesize that among Freud's lost metapsychological papers from 1915–17 there was an essay dealing explicitly with sublimation.

(7.) For instance, as late as his *New Introductory Lectures on Psychoanalysis* from 1933 Freud writes of psychoanalysis as a depth psychology that 'is quite unfit to construct a *Weltanshauung* of its own; it must accept the scientific one', and later in the same work writes 'science is content to investigate and to establish facts' (Freud 1933: 158, 162). While it may be a fact that in society X walking only on the cracks in the pavement is considered not socially valuable, the judgement that it not valuable and hence is a neurotic symptom rather than a sublimation involves an endorsement of the norms of that society and hence is beyond the purview of the mere establishing of facts. Alternatively one is left with the difficult relativist position that all that one can say as a statement of fact is that relative to the norms of society X it is sublimation and not a neurotic symptom. There is no evidence the Freud endorsed this kind of relativism, even if many post-Freudians would be happy to endorse it.

(8.) In *Beyond the Pleasure Principle* Freud gives perhaps his strongest formulation of the claim that sublimations do not involve the removal of repressions: 'The repressed instinct never ceases to strive for complete satisfaction ... No substitute or reaction formations and no sublimations will suffice to remove the repressed instinct's persisting tension' (Freud 1920: 42).

(9.) The first known references by Freud to sublimation occur in letters and drafts of letters to Fliess dating from 1897 (Freud 1897: 247). The first reference to sublimation in works published by Freud occurs in the Dora case (Freud 1905b: 50) and in the *Three Essays on Sexuality*, both published in 1905.

(10.) There is a good deal of literature devoted to expounding the relationship between Freud and Nietzsche. Unfortunately, much of it concerns the question of to what degree Nietzsche anticipated Freud and to what degree Freud did or did not cover up influences from Nietzsche. By and large, literature governed by these concerns tends to do little to illuminate the work of either Freud or Nietzsche. Perhaps the best, or, at least, most comprehensive, book of this questionable genre is Lehrer (1995). Less helpful is Assoun (2000). A most succinct example of this genre is A. H. Chapman and Mirian Chapman-Santana (1995).

(11.) Schopenhauer was more prone to speak of wills (*Willen*) rather than drives (*Triebe*). How much this is merely a terminological difference is a difficult question to answer. For Schopenhauer, beyond individual time-located willings, there is, notoriously, the transcendental notion of *Wille*, for which Nietzsche had no sympathy. However, Freud's notion of Eros may, arguably, be seen as a return to a more transcendental picture.

(12.) Note Nietzsche references are given not by page number, save for references to the KSA, but rather by section or aphorism number. For instance 'GM II 16' refers to section 16 of the second essay of Nietzsche's *The Genealogy of Morality*.

(13.) While Nietzsche never takes on anything approximating Freud's topological model of the id, ego, and superego (cf., for instance, *The Ego and the Id*, Freud 1923), Nietzsche countenances different modes of suppression and repression that have echoes in Freud's topological model. For instance, in GM I Nietzsche tells the story of slaves who initially suppress their drives principally because their reality principle tells them they are too weak to act on those drives in the face of the masters' oppression; later in GM Nietzsche tells how that suppression takes on a more moralized form leading to full-blown repression. Indeed, in sections 19–22 of GM II, Nietzsche tells how it is in reaction to our debts to our fathers, forefathers, and God that we develop what he calls a guilty conscience. The first kind of suppression is somewhat akin to suppression seated in the ego and the second kind is akin to repression seated in the superego (itself formed, according to Freud, through the internalization of god-like authoritarian father figures).

(14.) Reginster (1997) also argues that a kind of splitting of self is integral to *ressentiment*, claiming that 'ressentiment corrupts or dis-integrates the self' (301).

(15.) More generally Richardson (e.g. 1996: 25) gives an account of Nietzsche on sublimation that anticipates that given here. It is quite possible that this chapter's original inspiration comes from the excellent Richardson (1996) since, on reflection, that seems to be true of so much of the current author's work. Another source may have been Simon May (1999) who persuasively argues that sublimation, power, and form creation are Nietzsche's key criteria for value. For an amusing, but much more important, case of such 'anxiety of influence' I direct the reader to Bos (1992) which expounds the convoluted relationship between Nietzsche and Freud on the concept of the 'It' or 'Id' as mediated by the curious figure of George Groddeck, author of the somewhat notorious *Book of the It*.

(16.) It's worth noting that in section 205 from *Beyond Good and Evil*, quoted earlier, after describing those 'humans beings of late cultures' who can contain a multitude of drives fighting each other as 'weaker human beings', Nietzsche considers the case where:

a real mastery and subtlety in waging war against oneself, in other words, self control self-outwitting, has been inherited or cultivated then those magical, incomprehensible, and unfathomable ones arise … whose most beautiful expression is found … among artists perhaps in Leonardo da Vinci.

(17.) As Dina Emundts has pointed out to me, and as acknowledged earlier, such stifling, which can be short of full stifling, does not mean the weaker drive need find no form of expression. For instance, in Schreber's case his repressed homosexual drive found outlet in his paranoid fantasies and his transvestism. In the Christian slave, according to Nietzsche, his repression of his desire to have what the masters have is expressed in his proclaimed repudiation of the masters' attainments. The point is that in repression, as opposed to sublimation, those expressions typically take on the logic of opposites— Schreber's desire to be a female is expressed as his being forced by God to be a female against his will, the Christian slave's desire to have power over his masters is disguised as an alleged love of his enemies and his renunciation of the desire for power. So where repressed desires find expression they do so in a way that cannot be integrated into a coherent whole, they represent disintegration not an integration of the subject.

(18.) Note, it is not being claimed here that Nietzsche uses the term 'sublimation' with this exact meaning. While Nietzsche does occasionally use the term '*Sublimieriung*' and cognates in this writings (for instance, GM II 7) he never gives a thematized account of sublimation. The account of sublimation as unification given here is described as a Nietzschean account in that it is in line with many of his uses of that term and, more importantly, it is based on one of Nietzsche's central ideals of health. Even more importantly, this account serves to underwrite the very distinctions between repression and sublimation and between pathological symptoms and sublimations that Freud's account of sublimation fails to underwrite. This reading allows that in certain passages, for instance KSA 12: 254, Nietzsche uses the term '*Sublimieriung*' and cognates in ways more akin to Freud's actual usage and in ways that do not imply the notion of a united self. Schacht (1983) claims that in fact Nietzsche tends to use the term 'spiritualization (*Vergeistigung*) in the sense of sublimation. He further contrasts this to Nietzsche's term 'internalization' (*Verinnerlichung*). This accords with argument presented in the next section that Nietzsche's use of the term 'internalization' is better construed as repression rather than sublimation.

(19.) Similarly, Nietzsche sees the achievement of free will as something open to a limited few. For more on both these themes see Gemes (2006a).

(20.) The French term *Ressentiment* is expressly used by Nietzsche, most notably in GM, essay I.

(21.) For more on this see Gemes (2008) where affirming drives is not explicated in terms of taking a certain cognitive stance, endorsing some positive proposition about one's drives, but rather as fully expressing one's drives.

(22.) One needs to be a little careful here in avoiding the suggestion that Nietzsche favours some static permanent hierarchy among the drives. In fact, Nietzsche often emphasizes the need for a kind of agonal struggle between the drives. Agonal struggle is on the lines of a contest that develops, brings out, the best of the participants, rather than a struggle that leads to their evisceration. While I do not believe there is genuine conflict between the Nietzsche ideal of a unified self and the Nietzschean ideal of a self

engaged in agonal struggle, limitations of space prevent me from developing this point here. Suffice it for now to say that even a master drive, in order to fully develop itself, needs the conflict of robust challenges.

(23.) Kaufmann (1980) similarly lays great stress on the importance of the notion of sublimation for Nietzsche. In ch. 7, 'Morality and Sublimation', Kaufmann explicitly contrasts repression and sublimation as two methods for dealing with the chaos of impulses. Kaufmann places Nietzsche as a strong advocate of the second of these methods. Indeed Kaufmann, while noting that other modern philosophers including Goethe, Novalis, and Schopenhauer used the notion of sublimation, claims that is as Nietzsche who gave the notion the connotation of the transformation of sexual energy that it carries today (Cf. Kaufmann 1980: 219–20).

(24.) Since Nietzsche says little about the nature of master drives there is a certain amount of vagueness about what counts as a unified self for Nietzsche. With the element of social valuation absent there is the worry that, for instance, a reclusive obsessive stamp collector may count as a unified self. For such a person we may envisage a stamp collecting master drive. The answer here is that Nietzsche as a naturalist believes that as humans we come with a rich panoply of inherited drives—this allows that we may also in the course of acculturation acquire new drives. As a matter of empirical fact the reclusive stamp collector will not be a being who is giving expression to all his drives (for instance, drives to sociability, sexual drives, and aggressive drives).

(25.) In fact, some interpreters, including Heidegger (1979) and Adler (1928), insist that for Nietzsche there is one basic drive or force, namely, will to power, and all more specific drives are just modifications of this drive or force. But this is a difficult matter in Nietzsche interpretation that I can't here enter into.

(26.) Generally Freud emphasizes not unity but incompatibility between the demands of various sub-personal agencies, such as the ego, superego, and the id. This is also true of *The Ego and the Id* where he paints a picture of the ego being ever menaced by the conflicting demands of the reality principle, the id, and the superego. It is in this voice that he there labels Eros not as a uniter but as 'a mischief maker' (Freud 1923: 59).

(27.) I owe this observation about the relationship of internal objects to drives to a private correspondence from Bernard Reginster.

(28.) While, as noted earlier, most of Freud's texts on sublimation emphasize notions such as desexualization and social valuations, there are occasional texts in which he, like Nietzsche, emphasizes the unifying element in sublimations. In *The Ego and the Id* Freud tells us that sublimated energy may 'retain the main purpose of Eros—that of uniting and binding—as it helps towards establishing unity, or a tendency to unity' (Freud 1923: 45). However, more typically Freud labels Eros not as a uniter but as 'a mischief maker' (Freud 1923: 59). For more on this see Gemes (2009).

(29.) For more on this theme of self-estrangement see Gemes (2006b).

(30.) Golomb (1989: 71) similarly argues that for Nietzsche sublimation is to be preferred to repression because it allows for a certain kind of freedom, whereas repression tends to transform its bearers into passive, defensive creatures. Indeed Golomb reasonably concludes that while Nietzsche views the ordinary man as a creature of repression, he takes the extraordinary *Übermensh* to be a product of successful sublimation.

(31.) Nietzsche's relationship to Socrates is complex, involving elements of admiration and repudiation. The admiration involves, among other things, recognition of his success in framing a whole new moral universe and his courage in pursuing and facing death for that vision. But even in passages such as GS 340, where we see some of that admiration, Nietzsche characterizes Socrates as a 'monster' and as one who 'suffered life'.

(32.) There is much in this last paragraph that reflects the enormous influence of my colleague Andrew Huddleston on my thinking about Nietzsche. For more on this point see Huddleston (2014). Thanks are also due to Gudrun von Tevenar, Sebastian Gardner, Bernard Reginster, and to the participants of the St John's College Research Centre's Interdisciplinary Seminars in Psychoanalysis for helpful feedback on this chapter, and to Hanna Segal for discussions of her ideas about sublimation and how they differ from Nietzsche's. Initially I approached the topic of sublimation thinking that I could use Freud to clarify Nietzsche on sublimation. It was Sebastian Gardner who, on hearing a paper in which I gave my Nietzschean account of sublimation, helpfully suggested to me that perhaps it would be more advisable to use Nietzsche to clarify Freud on sublimation.

Ken Gemes

Department of Philosophy, Birkbeck College, University of London

Oxford Handbooks Online

Abstract and Keywords

The concept of wish-fulfilment as a substitutive mode of satisfying wish or desire was one of Freud's most important and singular psychoanalytic innovations. In his view dreams, neurotic symptoms, conscious and unconscious phantasies, delusions, hallucinations, jokes, much art, parapraxes, omnipotent thinking, illusions such as religion, the institutions of morality and social organization are all wish-fulfilling phenomena or attempts at wish-fulfilment. Although its remit is more circumscribed in contemporary psychoanalysis, wish-fulfilment can be seen to underlie such important conceptions as omnipotent phantasy, projective identification, actualization, and so on. This chapter exposes in detail Freud's singular innovation, relates it to some recent neuroscientific work, and shows how Freud's initial model of hallucinatory wish-fulfilment in dream and infant phantasy gratification, together with his conception of symbolism, can be extended to explain a range of symptoms as intentional behaviour, in line with Freud's ambitious claims for wish-fulfilment's remit.

Keywords: Freud, wish-fulfilment, symbolism, intentional action, self-solicitude

Ordinary Wish-Fulfilment and Freudian Wish-Fulfilment

When the seventeenth-century missionary Father Rageneau visited North America he found the Huron people distinguishing three causes of disease: natural causes; sorcery; unfulfilled wishes. Of the latter, some were known to the individual; others were not but could be revealed to him in his dreams. These pathogenic wishes could be concealed and then diviners were required to discover them. If the patient was mortally ill, the diviners would declare the objects of his wishes were impossible to obtain. If there were chances of recovery, a 'festival of dreams' was organized, a collection of valuables made among the tribe and presented to the patient amidst banqueting and public rejoicing. Patients not only recovered but sometimes emerged wealthy men.[1] Here is a perspicacious—almost 'Freudian'—diagnosis of the malady of frustrated (or ineluctable) desire, and a wonderful therapy. Here, pathogenic wishes, detected in dreams, are fulfilled in reality. It is important to see at the outset precisely how the ordinary conception of wishes being fulfilled instanced in this story differs from the topic of this chapter, Freudian wish-fulfilment (henceforth FW). In the ordinary conception of wish-fulfilment a person wishes for something; they get it; they know or believe that they get it and, as a result, cease wishing.[2] For their wishing to cease it is necessary that they at least believe that they get it: our cycles of desire and action are governed by information, not, or not typically, by exhaustion. And it is because they get it that they believe that they get it. Finally, it doesn't matter on this conception whether a wished-for object is attained by personal effort or by gift. Simplifying a little: where 'p' denotes any state of affairs:

> **(A)** In the *ordinary sense* of wish-fulfilment a wish that p is fulfilled only if:
>   **(i)** the wish is terminated: the agent ceases to wish that p;
>   **(ii)** the agent comes to believe that p;
>   **(iii)** p is actualized;
>   **(iv)** the wish is terminated *because* of the actualization of p.

The Freudian conception of wish-fulfilment is different. In the first edition of *The Interpretation of Dreams* (1900) Freud introduced a theory according to which dreams, neurotic symptoms, and conscious and unconscious phantasies, are wish-fulfilments. Eventually he concluded that positive psychotic symptoms—delusions, hallucinations—jokes and art, slips of the tongue, bungled actions, magical or omnipotent thinking, illusions such as religion, the institutions of morality and social organization were also wish-fulfilling or, recognizing that action can fall short of its object, attempts at wish-fulfilment.[3] The breadth of this array is striking: from the fictive representations of dream and phantasy to involuntary neurotic tics to the purposeful action sequences of religious observance, artistic creation, and character disorder. These phenomena, if indeed they are wish-fulfilling, are so in a manner patently different from the ordinary conception: in them, the objects of the wish are *not* attained, only a substitute or semblance is. FW is a

Subscriber: Tulane University; date: 16 December 2018

form of '*substitutive satisfaction*'. It is distinctive of the class that while conditions (i) and (ii) are necessary, conditions (iii) and (iv) are not. Characteristically in FW either the wished-for state of affairs is not actualized or, if it is, it has no causal role in the termination of the wish. The terminative role must therefore rest entirely with (ii) believing that p—or, perhaps, with experiential states functioning informationally and terminatively in the manner of belief. Moreover, with the possible exception of aesthetic appreciation, in all the cases described by Freud the agent initiates the wish-fulfilment. FW is something an agent does, either intentionally, subintentionally or by some expressive means.[4] It cannot be a gift. Thus, Freud's patient Dora sought by means of hysterical symptoms to separate her father from his mistress. Even if the symptoms were not manufactured intentionally by Dora, even if, say, they were dream-like expressions of desire or phantasy, they could still have succeeded in separating the lovers, and thus be wish-fulfilling in the Freudian sense. What would *not* count as FW is if, for example, a brick fell on the mistress's head.

On noting these features we may expect a spectrum of potential types of FW from simple non-intentional cases, to cases where intention plays some subsidiary role, to 'fully' intentional cases where FW is itself intended. To accommodate this spectrum we distinguish two conditions replacing conditions (iii) and (iv):

(B) For any wish that p, it is fulfilled in the manner of FW only if:

(i) the wish is terminated: the agent ceases to wish that *p* because:

(ii) the agent comes to believe (or, perhaps, experience) that *p*.

(v(a)) the agent initiates the wish-fulfilling process, in a sense that does not entail but does not exclude intention; or

(v(b)) in self-solicitous types, fulfilment of the wish can be truly described in some such way as 'A intentionally fulfils A's wish that *p*' or 'A intentionally gratifies (consoles, appeases) A'.

So, possible examples of the simple cases may be infantile hallucinatory wish-fulfilment and dreams, where representations caused by wishes in the primitive mental conditions Freud describes as primary process functioning achieve FW. Further along, representations of wished-for states of affairs may be generated intentionally, as in recalling fond memories or daydream, though not yet with express purpose of wish-fulfilment. In conditions of engrossment or reverie, however, vivid phantasy or recollection may ephemerally be mistaken for reality. At the furthest end of the spectrum are reflexive processes in which fictive representation, enactment, and manufactured or tendentiously selected 'evidence', is used intentionally for the purpose of fulfilling wishes. In the section on 'Self-solicitude' these reflexive, intentional forms of wish-fulfilment will be recognized as members of a class of self-solicitous activities, expressions of internal

object relations, that include self-consolation and care but also pernicious acts of self-appeasement and self-deception.

That is the conceptual outline of FW. It remains to show whether and how anything can coherently fall under the concept.

Freud's Basic Model

The concept of FW (though not of course articulated as stated earlier) is left undefined by Freud in several respects. He doesn't expressly distinguish FW from ordinary wish-fulfilment and the various conceptions of precursors such as the ancient oneiromancers[5]; and he often fails to distinguish between the *representation* of a wish being fulfilled, and the *fulfilment* of a wish (either as process or end state), though this distinction has significant consequences. Thus he writes of the famous Irma dream that its 'content was the fulfilment of a wish and its motive was a wish' (Freud 1900: 119) and of a 'delusion having as content the fulfilment of the wish' (Freud 1922: 226). But other passages go further: '*Dreams are … guardians of sleep which get rid of disturbances of sleep …* a dream does not simply give expression to a thought, but represents the wish-fulfilled as a hallucinatory experience … the dream does not simply reproduce the stimulus [the wish], but removes it, gets rid of it, deals with it, by means of a kind of experience' (Freud 1916–17: 129). Again: 'Symptoms serve for the patient's sexual satisfaction; they are a substitute for satisfaction of this kind …' (Freud 1916–17: 299). The idea that FW terminates or 'pacifies' (Hopkins 1995) wishes and provides substitutive satisfaction is in fact a theoretical necessity in Freud's scheme. Mere representation of wishes fulfilled fails to capture the specific functions assigned to dreams and symptoms: to preserve sleep, resolve conflict, circumvent realistic action. FW involves less than full satisfaction of a wish in the ordinary sense but more than mere representation: it involves, as well (usually temporary) termination of the wish.

Freud describes the simplest form of FW, the hallucinatory wish-fulfilment of the infant, in a famous passage:

> A hungry baby screams or kicks helplessly … the excitation arising from the internal need is not due to a force producing a *momentary* impact but to one which is in continuous operation. A change can only come about if in some way or other (in the case of the baby, through outside help) an 'experience of satisfaction' can be achieved which puts an end to the internal stimulus. An essential component of this experience of satisfaction is a particular perception (that of nourishment, in our example) the mnemic image of which remains associated thence forward with the memory trace of the excitation produced by the need. As a result of the link that has thus been established, next time this need arises a psychical impulse will at once emerge which will seek to re-cathect the mnemic image of the perception and to re-evoke the perception itself, that is to say, to re-

establish the situation of the original satisfaction. An impulse of this kind is what we call a wish; the re-appearance of the perception is the fulfilment of the wish; and the shortest path to the fulfilment of the wish is a path leading direct from the excitation produced by the need to a complete cathexis of the perception. Nothing prevents us from assuming that there was a primitive state of the psychical apparatus in which this path was actually traversed, that is, in which wishing ended in hallucinating. Thus, the aim of the first psychical activity was to produce a 'perceptual identity'—a repetition of the perception which was linked with the satisfaction of the need (1900: 565–6; cf. 1950: 317–19, 332–3).

First, there is the perception of a real experience of satisfaction achieved through 'outside help'. Subsequently when an internal need (or stimulus) gives rise to a wish (or psychical impulse) and their—presumably it is both the need and wish—immediate satisfaction is blocked, as it is in the helpless unattended infant and sleeper, a topographical regression, a backward movement of psychic energy to the perceptual system of the mental apparatus, ensues and a reactivation of a memory of satisfaction as a hallucinatory experience provides a 'perceptual identity' with 'an essential component' of the original experience of satisfaction. A memory of a previously satisfying experience is 'mistaken' for a perception of the original experience and it is for the infant *as if* its wish has been satisfied: the wish is terminated. This outcome is characteristic of the primary process mode of mental functioning whose key characteristics include mobility of cathexis and '*the replacement of external by psychical reality*' (Freud 1915b: 187). The hallucinatory or fictive experience of satisfaction (henceforth EOS) that terminates the wish does not terminate, but occludes, the endogenous stimuli that gives rise to the wish. Hunger gives rise to a wish to be fed; the infant hallucinates being fed and this pacifies the wish but not the hunger. Clearly pacification will be the more effective the more remote a wish is from the needs which give rise to it.

We are to imagine a similar process occurring in the rest of the wish-fulfilling series: dream, symptom, art, and so on. In essence, this is an informational conception of FW: wishing is terminated when a belief or, perhaps, an information-laden experience, registers that the wished-for state of affairs has been actualized. In the early *Project* (1950), Freud exposed an underlying energic or mechanical model on which the achievement of an EOS has several far-reaching neuronal consequences. One of them is a 'lasting discharge' which brings the 'urgency which had produced unpleasure in *w* [consciousness]' to an end, thereby reducing the psychical excitation of free (unbound) energy in accord with the overriding principle of inertia (or constancy) (Freud 1950: 318). The reduction (or elimination) of psychical excitation as the governing principle of mind remained a constant in Freud's metapsychological formulations. It is notable, however, even in this early work, that the energic conceptions ride on Intentional or common-sense-psychological ones; and this is of a piece with Freud's gradual realization that these idioms were indispensible for describing mind (cf. Freud 1895: 160–1). He realized, for example, that the regressive revival of a memory-percept in dream is by itself insufficient for the efficacy of the EOS; for the EOS to be efficacious the memory percept must, as well, 'meet with belief' (Freud 1950: 322–3, 339). Thus 'we could well imagine

the dream-work penetrating to mnemic images of this [regressive] kind, making conscious to us what was previously unconscious, and holding up to us a wishful phantasy which rouses our longing, but which we should not regard as a real fulfilment of the wish. Hallucination must therefore be something more than the regressive revival of mnemic images that are themselves Ucs' (Freud 1917: 231). So how do the mnemic images 'meet with' belief? In the state of sleep, Freud says, attention cathexis are withdrawn from 'the system Cs. as well as the other systems, the Pcs. and the Ucs., in so far as the cathexis in them obey the wish to sleep. With the system Cs. thus uncathected, the possibility of reality testing is abandoned; and the excitations which, independently of the state of sleep, have entered on the path of regression will find the path clear as far as the system Cs. where they will count as undisputed reality' (Freud 1917: 234). In the conditions Freud describes ideas or representations of wished-for states of affairs brook no contradiction and function as wish-fulfilling representations: psychical reality is mistaken for external reality. One might say that from the perspective of the dreamer the fictive EOS or hallucination functions as irrecusable evidence for wish-fulfilling beliefs. In such conditions, we may say, a person is *engrossed*.[6]

That is the basic model. We have noted that FW compasses as well as dreams and hallucination, neurotic and some psychotic symptoms, religious devotion, the making and experience of art, and so forth, but with the exception of 'symbolic wish fulfilment', understood in a wide sense, Freud provides little guidance on how the basic model can be extended to accommodate these other putatively wish-fulfilling phenomena. The latter generally involve intentional action, and it does appear to have been Freud's view, and I think it is the correct view, that some forms of FW are brought about intentionally, either as adventitious consequences of intentional acts or as directly intended i.e. where FW is itself intended in an act of self-solicitude. The basic model is evidently based on an incipient neuroscience that is theoretically meagre even for explaining dreams and hallucination, and especially so for the more complex cases of FW. The kernel of Freud's model can be extended, however, with considerable plausibility, as following sections show.

Representation and Symbolism

Apart from the transparent dreams of young children, few dreams, let alone symptoms and their kin, have conscious content that is manifestly wish-fulfilling. It is Freudian doctrine that these phenomena are mostly *disguised* wish-fulfilments, partly as an effect of unconscious primary processes and partly because of defensive activity to protect consciousness from disruptive stimuli. Freud distinguishes latent from manifest content in dream and symptom. In the case of dreams, the latent content or 'dream thoughts' consist of recent day residues interwoven with older aroused emotion-laden memories and wishes, and this content is transformed into the manifest, consciously experienced dream—the fictive experience of satisfaction or EOS (or cluster of EOSs)—by the

processes of the 'dreamwork': condensation, displacement, conditions of representability, secondary revision, and symbolism. So, famously, dreams require interpretation, and that ultimately involves uncovering the group of wish-fulfilments implicit in the dream: 'Dreams frequently seem to have more than one meaning. Not only … may they include several wishfulfilments one alongside the other; but a succession of meanings or wishfufilments may be superimposed on one another, the bottom one being the fulfilment of a wish dating from earliest childhood' (Freud 1900: 219). Of his 'specimen' dream of Irma's injection Freud writes:

> The dream fulfilled certain wishes which were started in me by the events of the previous evening (the news given me by Otto [that Freud's patient Irma was still unwell] and my writing of the case history). The conclusion of the dream, that is to say, was that I was not responsible for the persistence of Irma's pain, but that Otto was. Otto had in fact annoyed me by his remarks about Irma's incomplete cure, and the dream gave me my revenge by throwing the reproach back on to him … *I* was not to blame for Irma's pains, since she herself was to blame for them by refusing to accept my solution. *I* was not concerned with Irma's pains, since they were of an organic nature … Irma's pains had been caused by Otto giving her an incautious injection of an unsuitable drug—a thing I should never have done
>
> (Freud 1900: 118–19).

Freud's detailed analysis is too ramified to present here (see Hopkins 2015: 71–85; also Hopkins this volume) but the wish-fulfilments contained in the dream may be roughly divided into relatively obvious ones such as Irma's pains being organically caused, where the dream text practically wears the EOS on its sleeve ('Dr. M. … confirmed it … "There is no doubt it is an infection"'); and those that are disguised by the dreamwork. The 'incautious injection' with which Otto is rebuked leads Freud to partially repressed painful memories. Years earlier he had injected his patient Mathilde with a substance then considered harmless, resulting in her death; and on his advice Freud's friend Ernst began using cocaine for the relief of intractable pain, and by misuse quickly succumbed to it. So Otto with his thoughtless and probably dirty syringe is blamed not only for Irma's condition but also for the deaths of Mathilda and Ernst. If Otto is to blame then Freud is not. If FW is to succeed then it must be for the dreaming Freud *as if* Otto was responsible for the earlier deaths even though there is no conscious EOS to that effect. How does the dream do it?

At first pass, FW is effected by elements of the manifest dream representing multiple elements in the latent dream thoughts, utilizing condensation, displacement, and symbolism. Condensation leads chiefly to the formation of 'collective and composite images' of figures or actions. So we are told that Irma 'stood for' Freud's eldest daughter, his wife as well as Mathilda, and that she acquired a series of still 'other meanings' and 'allusions': 'Irma became the representative of all these figures which had been sacrificed to the work of condensation, since I passed over to *her*, point by point, everything that reminded me of *them*' (Freud 1900: 293). But how precisely does such representation by

means of condensation facilitate FW? Freud's unsatisfactory answer is that in condensation 'ideas are formed which are endowed with great intensity' and 'the intensities of the individual ideas become capable of discharge *en bloc* and pass over from one idea to another …' (Freud 1900: 595). This is not helpful, of course, because we have no idea how to relate *meaning* or representation to *psychical intensity*, or the facilitations effected by relations of meaning to discharge. Moreover, it turns on the awkward proposition that in condensation one element of a composite figure stands for other elements of it. Whatever we come to make of these problems we can see why Freud also considers displacement as a form of representation, for it too involves a transfer of 'psychical intensity' or 'value' (distinguished from 'sensory intensity'), in which an element of low value comes to stand for one with high value.

Condensation and displacement do not exhaust the relation of representation. Freud eventually recognized an exclusive type of representation—symbolism—attended with conditions distinguishing it from the other types. Symbols in this distinctively psychoanalytic sense are semantically opaque to the subject because their meaning or referent is either repressed or innately and unreflectively given, conditions not imposed on the other processes (Freud 1901a: 683). This is the sense in which a snake can symbolize a penis without condensation, or even displacement. Symbolism in this strict sense need not depend on these processes of the dreamwork.[7]

Putting aside Freud's unhelpful explanation in terms of energic discharge, can we explain how representation, either as symbolism (strictly) or in the awkward manner Freud attributes to condensation and displacement, facilitates and extends FW? Two considered theses recommend themselves. In regard to symbolism, Petocz (1999: 233) argued that the wish-fulfilling efficacy of the symbol rests on an unconscious belief in the identity of symbol and symbolized. As a consequence 'the symbol is mistaken for the symbolised and treated as if it were the symbolised' leading, 'for reasons not properly understood', to a kind of 'gratification which is not as complete as would be the gratification obtained from the satisfaction, via primary objects and activities, of the unopposed instinctual impulse' (Petocz 1999: 233). Such a belief in identity cannot be excluded (see the section on 'Intention in Wish-Fulfilment'), but it does not seem to be an indispensible condition for symbolic FW. It may be that an operation on the symbol is simply taken to be, is apprehended as, an operation on its referent, without belief in their identity.[8] The sorcerer who sticks pins in an effigy need not believe the effigy *is* his enemy. The supposed causal relations between effigy and referent are supported by an immense raft of beliefs and practices; there appear to be many different ways in which symbolic relations are established, and that may be true of unconscious symbolism also.

That reflection suggests the approach in Hopkins (2000; 2016) which bundles condensatory and (strictly) symbolic forms of representation under the head of symbolic cognition. It draws attention to the pervasiveness of symbolic and metaphoric associations informing all thought and action, naturally established over the course of life, so that 'present significant figures and situations partly inherit their meanings from past ones, to which they are unconsciously mapped' (Hopkins 2000: 12). So, given

Freud's fixed association between Irma and Mathilda, in representing Otto as giving a toxic injection to Irma, Freud at the same time pacifies his wish not to have been responsible for the death of his earlier patient. The pacification of the contemporary desire 'went together' with that of the deeper, more painful one (Hopkins 2000: 12). It would seem that for FW to succeed on either of these views—or on related views of unconscious representation—there must be *some* occurrence of unconscious understanding or apprehension underwriting them. Whether, as on Petocz's account, an unconscious belief in the identity of Irma and Mathilda entails that an operation on the one is experienced as if it were an operation on the other; or whether, as on Hopkins's account, representations of Irma and Mathilda become condensed and interchangeable through association; the pacification of Freud's wish that it was Otto (and not Freud) who had killed Mathilda would require that it be for him *as if* Otto had killed Mathilda. That would seem to entail *the acquisition of a belief that Otto killed Mathilda*—for there *is* no EOS whose content is Otto killing Mathilda. The passage from the EOS—the tableau which has Otto (almost) injecting Irma—to the content of Otto killing Mathilda requires a process of thought arriving at an unconscious understanding or belief that it is so.[9] So it would appear from the common-sense-psychological perspective. The symbolic associations between Irma and Mathilda may perhaps be best understood in sub-personal terms, but the wish-fulfilling belief about Otto killing Mathilda cannot be.

A Neuroscientific Model

The neuroscience upon which Freud drew was in its infancy and his conception of psychical energy was a placeholder waiting on the deliverance of more advanced science. In a series of papers Hopkins (2012, 2015, 2016, and this volume) has integrated fundamental aspects of Freudian theory with the recent 'free energy' neuroscience developed by Karl Friston and his colleagues, one of several neuroscientific models currently in contention. Friston's information-theoretic conception of free energy (FE) is not Freud's but, as Hopkins points out, they both assign FE the same overall functional role in the brain. In Freud's work (1950; 1911) and under Friston's (2010) FE principle the brain seeks to minimize FE aroused by sensory and interoceptive impingements. I summarize some relevant background to Hopkins's synthesis before focusing on the role the experience of satisfaction (EOS) plays in the account of wish-fulfilment that emerges from this neuroscientific model.

In the terms of this account (which bears similarity to Freud's notion of primary process and Kleinian unconscious phantasy), the brain comes equipped with an innate model of the world and a 'virtual reality generator' (Hobson, Hong, and Friston 2014), which generates predictions about sensory and interoceptive impingements and, by testing them in awake states, constructs increasingly veridical models of the world. A basic claim of FE neuroscience is that 'conscious perceptual experience is a form of explanatory hypothesis, that at once unifies, predicts, and inhibits the impingements that cause it,

thereby minimizing the FE associated with them' (Hopkins 2016: 3). The FE of a model is the difference between its complexity and its accuracy. *Accuracy* (roughly) is a measure of how well the model is predicting the data. *Complexity* reflects the number of parameters or potential trajectories for action composing the model and the extent to which they are modified in coming to predict data. Increase in complexity (without increase in accuracy) increases free energy. The imperative to minimize free energy therefore entails reduction in complexity, rendering the models more efficient in making predictions. During the day, it is argued, the huge influx of sensory input increases prediction error and, so, complexity. During sleep, with abatement in sensory input, and accuracy irrelevant, the brain or 'generative model' is freed to reduce complexity, enabling it to generate more efficient predictions in waking. It does this by, amongst other things, manufacturing in REM dream fictive or counterfactual virtual reality—various EOS—which inhibit realistic movement trajectories or parameters, including those associated with emotional and motivational conflict.

Hopkins makes the important suggestion that emotional and motivational conflict (too many or conflicting trajectories of action) and trauma (required emotional adjustments cannot be integrated) may plausibly be viewed on this model as sources generative of neurocomputational complexity. So in sleep, for example, the most significant nocturnal disruptions arise from recent emotion-laden memories of the sort Freud described as 'day residues' and those remotely linked and often disturbing emotions and wishes which are aroused and then condensed or consolidated with the more recent ones. Given the role of the brain as conceived above we can see how dreams and symptoms generic to phantasy or hallucinatory experience, which are viewed in psychoanalysis as ultimately products of conflict and trauma, may in large measure be productions functioning to reduce complexity. For if we now take the arousing materials of the dream as input, then, in accord with the basic claim of FE neuroscience, 'the fictive perceptual experiences of dreaming [i.e. experiences of satisfaction] could serve to unify, predict, and inhibit these arousals, thereby minimizing FE (as emotional conflict/complexity) in a way closely analogous to the realistic perceptual experiences of waking' (Hopkins 2016: 4). So on this view, the brain calculates the optimal motor trajectory for reducing free energy (on multiple interacting neurocomputational levels) by predicting *sensory trajectories* for courses of action. The EOS is then the culminating part of the sensory trajectory represented as conscious experience. But the EOS is not just the phenomenal end product of a sub-personal process; it is posited as having a causal role. In dream and symptom fictive experiences of satisfaction are figured as mitigating conflict, thereby reducing complexity and, so, FE, by eliminating incipient or conflicting trajectories or possibilities of action.

Precisely how does a fictive EOS reduce or eliminate conflicting parameters or potential trajectories? In essence this is a reformulation of the question of how the EOS is efficacious in FW. Recall the discussion of Freud's dream of Irma's injection. On Hopkins's view we have a sub-personal generative model acting in response to the dream

arousals 'inhibiting (pacifying) them by creating Freud's fictive experience of innocence and vindication, and so reducing the emotional complexity (adversity, conflict, and potential to cause depressive trauma) of the whole family of memories under reconsolidation' (Hopkins 2016: 14; this volume). The dream, he says:

> introduced an *ideal alternative counterfactual model* which perforce realized *an ideal fictive sensorimotor trajectory* by which, and in accord with Freud's priors of the previous day, the adverse emotional situation precipitated in him by Otto's unwelcome remark could be restored to a personally satisfactory (and free-energy minimal) equilibrium. Thus Freud could wake in the morning free from the self-justificatory and depressive trend in which he had spent the day and evening after Otto's remark ...' (2016: 15).

Now these passages hint at a number of perplexities that arise, I think, from the unresolved conjunction of sub-personal and personal perspectives. We can start with contrasting the view of dreams being exposed here with Freud's. According to Freud, the dream produces a transient EOS that pacifies disruptive wishes in a hallucinatory manner and guards sleep; it does not effect a lasting change in mind. It is perhaps in this vein that Hopkins wrote of the dream 'as a kind of *perfect and all-encompassing miniature hallucination*, in which the deluded dreaming subject utterly obliterates both what is happening in his mind and how things are in the world ...' (Hopkins 2012: 258). Here it is clearly engrossment in the EOS—the all-encompassing hallucination—that represses or at least pacifies parameters (conflicting trajectories, wishes etc.) inconsistent with the brain's dominant model. Wishes are pacified but they are not terminated. On the later account (Hopkins 2016; this volume), incompatible parameters are terminated or readjusted to achieve reduction in complexity. The dreamer receives not just a transient substitutive satisfaction enabling him to continue sleeping: his mind is permanently modified; thus, as Hopkins says, Freud awakes free from his depressive, self-accusatory mood to pursue the day with enhanced creativity. Perhaps that is so, or partly so. But then the generative model's work of emotional consolidation, conflict elimination, and inhibition as described by Hopkins at a sub-personal level seems to leave open the causal role, if any, of the Intentional and phenomenal properties of the EOS in complexity reduction. A situation where the generative model does indeed create an EOS, and the EOS does its work of complexity reduction, but entirely by virtue of its neural properties, not its Intentional and phenomenal properties, seems quite conceivable. In that case the *experience* of that complex tableau of EOS dreamt by Freud would seem to be epiphenomenal, causally inefficacious.[10]

If that is possible then there would appear to be ambiguity, if not incoherence, in this neuroscientific account of wish-fulfilment. But even if that is not right and there is no contradiction between what the sub-personal mechanism yields and the alleged Intentional causality of the EOS, the account would still seem to be deficient. As I argued in preceding sections, for a dream-wish to be fulfilled it must be for the dreamer *as if* the wish's object had been attained and *that* in most cases—given the absence of a

corresponding EOS—requires the *belief* that the object has been attained. It does not seem to me that the account of FW considered in this section satisfies that condition.

Whether or not that is the correct understanding of it, both the Freudian basic model and this more sophisticated neuroscientific one appear limited in accounting for the full range of FW as Freud envisaged it. Hopkins's thesis is directed at explaining dream, phantasy, neurotic and psychotic thought, and delusion. It is unclear (to this author, anyway) how it is to be elaborated to explain wish-fulfilling action, the extended ordered sequences of action constituting personality disorder, religious devotion, and aesthetic experience. The next two sections make an attempt to explain some of these phenomena along Freudian lines, retaining their presumptive wish-fulfilling character. But other, and perhaps more attractive, alternatives may present themselves: (i) to abandon Freudian wish-fulfilment as the fitting mode of explanation for these phenomena or (ii) to abandon altogether the attempt to provide Intentional explanation for the pathological developments Freud (and others) charted.[11]

Intention in Wish-Fulfilment

We said that the Freudian account of wish-fulfilment is intended to compass a wide array of activities including those that, on the face of it, are deliberate and intentional. The therapeutic imperative to formulate aetiologies for his clinical cases led Freud to hold: that neurotic symptoms, including enactments (obsessional routines, for example), are caused by unconscious wishes, reasons, and intentions; that these motives are not categorically different from corresponding conscious motives; and that they cause actions in much the same ways as conscious motives do.[12] That symptoms and their kin may have unconscious causes is relatively uncontroversial; the stronger contention that they may be produced intentionally, indeed as the end of practical reasoning, is problematic, and few have followed Freud so far. The view is counter-intuitive: symptoms certainly *appear* to be unintended, involuntary, and unpleasant. Again, *intention*, of all motivational concepts, has the closest affinity with conscious deliberation: 'unconscious intention' looks like an oxymoron. *Unconscious intentional agency* entails unconscious practical reasoning, acting, or resolving to act upon practical syllogisms—schemas whose main premise is a desire, minor premise an instrumental belief, and conclusion an intention or action. But as we shall soon see that is a taxing entailment. Moreover, given that in most typical behavioural symptoms conscious intentional agency remains intact, the operation of countervailing unconscious intentional agency would at least in some circumstances entail an apparently untenable partition or dissociation of self into quasi-independent, self-like centres of agency (Pataki 2003: 162–76).

Because of these difficulties most of the philosophers who reject partitive conceptions of mind as incoherent pursue non-intentional analyses of dream, wish-fulfilling phantasy and symptom; and largely ignore the role of FW in art, religion, and social action. My view, however, is that the concept of unconscious intentional agency is coherent and indeed essential for the explanation of the causal structure of many typical disorders and types of activity. To convincingly establish that view requires a larger canvas than I have here, but I can begin to sketch the main lines of its development.[13]

In Lecture 17 of the *Introductory Lectures* Freud discusses an obsessional symptom in a nineteen-year-old girl:

> [T]he most important stipulation related to the bed itself. The pillow at the top end of the bed must not touch the wooden back of the bedstead … The eiderdown had to be shaken before being laid on the bed so that its bottom end became very thick; afterwards, however she never failed to even out this accumulation of feathers by pressing them apart … She found out the central meaning of the bed ceremonial one day when she suddenly understood the meaning of the rule that the pillow must not touch the back of the bedstead. The pillow, she said, had always been a woman to her and the upright wooden back a man … If the pillow was a woman, then the shaking of the eiderdown till all the feathers were at the bottom and caused a swelling there had a sense as well. It meant making a woman pregnant; but she never failed to smooth away the pregnancy again, for she had for years been afraid that her parents' intercourse would result in another child and so present her with a competitor (1916–17: 265–8).

Freud understood this nightly ritual as the girl's wish-fulfilling attempt to prevent her parents' sexual intercourse and to resist anxieties connected with the birth of a competitor. The performance certainly presents as a sequence of intentional actions, but initially the girl could provide no reasons for performing them and found her own actions unintelligible. Before sleep anxious thoughts presumably assailed her: 'if I want to get some sleep then I had better separate the pillow from the bedstead etc.—I can't help doing that anyway'. But such thoughts do little to illuminate the causal structure of the symptom. Her performance is involuntary: against her conscious will. She doesn't know why she does what she does; she seems a mere spectator of her own actions. In not knowing her efficacious motives she doesn't consciously know *what* she is doing, under the descriptions that would be revelatory of why she behaves as she does. If we follow Freud's principal interpretation then as a preliminary attempt we can construct a practical syllogism for her behaviour:

> Girl wishes to separate mother and father.

> Girl believes that by separating bedstead and pillow she will separate mother and father.

> --

> Girl separates bedstead and pillow.

The desire expressed in the major premise and the belief expressed in the minor premise are unconscious but if her performance accords with this syllogism it is plainly intentional. Then given her belief and action she will have achieved the Oedipal wish expressed in the major premise in the manner of FW: it will be for her *as if* she had separated her parents: her wish is pacified, anxiety is allayed, she sleeps. The glaring difficulty is that the minor premise expresses a 'mad belief': who could hold such irrational or bizarre beliefs, or act on them?[14]

We can in fact construct a number of practical syllogisms for the performance by considering again the issue of representation and symbolism. On the face of it the girl appears to unconsciously believe that if she separates the bedstead and the pillow she will have separated her parents. Believing this proposition is not the same as, nor does it presuppose, believing that her parents were identical with the furnishings (although the latter is not a belief that can be ruled out *tout court*). The sorcerer who sticks pins into an effigy doesn't believe that the effigy is his enemy. The distinction here is between one thing *symbolizing* another and that thing being *believed* to be identical with the other. And both of these can be distinguished from a third, where one thing is *equated* with another: where there is something like condensation or what Segal (1991) calls symbolic equation in which the indiscernibility between symbol and thing symbolized is such that an operation upon one is experienced as an operation on the other. So there are the following possibilities: if the girl believes unconsciously that the bedstead is her father and the pillow her mother then, from the fact that she separates bedstead from pillow she could unconsciously *infer* that she has separated her parents. In the case of symbolism, what the girl does effects—somehow—to separate her parents, as she apprehends it unconsciously. In the case of symbolic equation she again does not infer, she simply apprehends the performance as a separation of parents since operating on the furnishings is experienced by her as emotionally the same as operating on parents.

The practical syllogism set out above is an example using the sorcery-like belief 'if I separate the furnishings I will separate parents'. The following syllogism incorporates a symbolic equation:

> Girl wishes to prevent the conception of a baby-competitor.

> Girl believes that by keeping her father = bedstead and her mother = pillow apart she could prevent that conception.

> --

> Girl separates bedstead and pillow.

In each case (and also in the case of symbolism strictly), the girl performs an act rationalized and caused by her wish and the relevant belief, thereby intentionally fulfilling her Oedipal wishes in the manner of FW. It seems that each of these practical syllogisms articulates a possible instance of unconscious deliberation, and explains the ritual as an instance of FW. (Possible, that is, given certain psychoanalytic and philosophical premises undefended here.) But the case is underdescribed and we need not settle on a particular explanation. Moreover, there are still other possible schemas that do not trade on mad beliefs, symbolism, or symbolic equations, yet satisfy the requirements for FW as self-solicitude expressed in the opening section of this chapter. But before turning to those we are bound to examine the most cogent of the non-intentional alternatives.

It is objected that the rational reconstructions of the girl's behaviour considered here over-intellectualize a simpler situation. One important and influential approach was early advanced by Hopkins (1982, 1988). Hopkins placed FW at the centre of the common-sense-psychological explanation in the psychoanalytic domain but explicitly contrasted FW with intentional action. FW was conceived subintentionally, as supervening on imaginative activity animated by desire. 'In rational action', he wrote, 'motives produce willed intentions and real actions aimed at satisfaction. Here [in dreams] they produce wishes and mere representations of satisfaction, on the pattern of wishful imagining' (Hopkins 1988: 41). The subintentional mode of wishful imagining was also enlisted to explain symptoms such as the Ratman's obsessive thoughts of torture (Hopkins 1982: xxi–xxvii).

We may concede that, at the common-sense-psychological level, subintentional causation suffices for explaining the motivation of dreams, some phantasying, and symptoms such as the Ratman's obsessive thoughts. But it cannot fully explain more complex symptoms such as the Ratman's obsessive *actions*, the compulsive bedstead girl's rituals, or phenomena like wish-fulfilling religious practices. Suppose we try to explain the girl's performance as a product of wishing expressing itself *directly*, subintentionally, in action, as if it were dream, imagination, or phantasy? In that case no element of her action can be intentional: being caused subintentionally excludes a priori any action from having a reason or intention among its causal conditions. So if the performance *is* the phantasy, and the phantasy is the product of 'an exercise of will of a kind prior to that in intentional action' (Hopkins 1982: xxiv), then it cannot be intentional. But the performance has the hallmarks of being a sequence of intentional actions, features which this account is unable to explain.

A better interpretation of the proposal is this: the girl, agitated and *distrait*, separates bedstead and pillow and, given the (broadly, including condensatory) symbolic relations between parents and furnishings, she comes to *imagine* or phantasize unconsciously that she has separated her parents. Her actions are intentional, her wish-fulfilling imaginings are not (Hopkins 1982: xxv). On this attractive account the imagining and the consequent fulfilment of her wish supervene on an intentional performance that has acquired contingent symbolic meaning. Well, how exactly does the performance fulfil the girl's unconscious wish to separate her parents? On the view that a symptom is wish-fulfilling merely by virtue of being a representation of its originary wish-fulfilled the answer appears straightforward. But that view of FW has been rejected (in the section on 'Freud's Basic Model'): for representation alone, we saw, does not suffice for pacification and therefore for FW. The representation will pacify only if it is constitutive of the circumstance that *it is for her as if* her parents have been separated—a mental state that would seem to involve at least an unconscious belief or understanding that they had been separated. Without supplementation by some notion of unconscious understanding it is quite mysterious how the performance could acquire its pacific powers. *With* such supplementation we do arrive at FW: the girl does something: the act is imagined or understood unconsciously so as to generate a wish-fulfilling belief. But even so the account remains deficient. The wish-fulfilment is adventitious: the girl does something

which happens in the circumstances to pacify an unconscious wish. We do not yet have an understanding of *why she did what she did*. We seem driven to suppose that she did precisely what she did because she understood unconsciously that that very sequence of actions would achieve her aim. If not that, then two counter-alternatives suggest themselves: either she hit upon the wish-fulfilling performance by chance or her performance is determined fixedly by the symbolic associations established in the course of her life. Chance, I think we can exclude. On the other option we have it that given her fixed associations it is natural that she should so act; just as it is natural for the old bachelor to collect snuffboxes, given the unconscious meaning they have for him (cf. Hopkins 2000: 12). So, how do unconscious symbolic associations determine her performance? It is hard to say, but it seems that it cannot be *merely* through a fixed sub-personal association: for that would deprive the performance of its intentional character. On the other hand, if the symbolic associations do operate through the prism of unconscious understanding that would concede the present argument that she knew what the consequence of her performance could be.

However that uncertain issue is resolved, the subintentional approaches confront a further difficulty: the fact that pathological behaviour of very much greater complexity seems also to instance intentional FW. Some complex pathological enactments, for example wish-fulfilling character disorders, unfold over a lifetime. In such cases a series of otherwise unconnected acts, in diverse circumstances, strung out temporally, combine to achieve the one (say, self-defeating) purpose. A thousand acts pointing in the same direction. It seems inconceivable how an array of subintentional mechanisms, set of fixed symbolic relations, or chance actions could be strung together in the service of the one unified purpose: such behaviour seems to manifest resolution and plan. If such clinical cases cannot be explained non-intentionally then self-solicitous FW may offer some promise.

Self-Solicitude

In the first section of this chapter provision was made for cases of wish-fulfilment which fall under the heading of self-solicitude, cases in which the self takes itself as an object—mostly to be cared for, but allowing for the pathological permutations of such reflexive relationships that psychoanalysis has uncovered. We must now consider whether there are in fact psychological configurations that exemplify the self-solicitous forms of FW.

Granting a partitive conception of mind in which one part (B) is in solicitous relation to another, 'object' part (A), we can generate valid practical syllogisms for B:

B wants to console (protect, deceive etc.) A

B believes that by separating bedstead and pillow A will be consoled (protected, deceived etc.)

B separates bedstead and pillow.

Here B's belief is not mad; it is the kind of belief a concerned mother wanting to console or pacify a child may have. Construing the self as a unity of parts (in the manner indicated below) we recognize an expression of self-solicitude that was frequently described by Freud and other pioneer object-relations theorists.[15] Freud, for example, writes of the ego's protection and care-taking of the id (1940: 197–9) and of the superego's solicitude for the ego: 'To the ego … living means the same as being loved—being loved by the superego … The superego fulfils the same function of protecting and saving that was fulfilled in earlier days by the father …' (1923: 58); and 'if the superego does try to comfort the ego by humour and to protect it from suffering, this does not conflict with its derivation from the parental function' (1927: 166). That such self-solicitous relations are possible is a consequence, in Freud's view, of identification with objects resulting in structural differentiation shaping several features of the self[16]: the capacity to experience itself as divided; its capacity to divide, indeed for 'a real split' to occur between the agencies (1926: 97); and the capacity of parts of the self to then take other parts as objects. In this vein, Freud writes of the superego:

> The ego can take itself as object, can treat itself like other objects, can observe itself, criticize itself, and do Heaven knows what with itself. In this, one part of the ego is setting itself over against the rest. So the ego can be split; it splits itself during a number of its functions—temporarily at least. Its parts can come together again afterwards (1933: 58).

Freud's conception of the self is clearly partitive: the agencies are conceived as capable of entering into a range of mutual 'personal' relations, for example sadomasochistic relationships; they can function autonomously with their own perspectives, motives, and capacity for intentional action; they have a degree of mutual opacity. Despite these conditions Freud seems to have enjoined some kind of underlying unity in the self.[17] How might such extreme conditions of psychic stress be exemplified clinically? A patient:

> would rave against girl children and in fantasy would describe how she would crush a girl child if she had one, and would then fall to punching herself (which perpetuated the beatings her mother gave her). One day I said to her, 'You must be terrified being hit like that.' She stopped and stared and said, 'I'm not being hit. I'm the one that's doing the hitting'

> (Guntrip 1968: l9l).

This person certainly seems radically dissociated, a 'real split' having occurred (see footnote 15). There appear to be at least two internal perspectives: in the part doing the punching and in the part being punched. The parts appear to be mutually alienated: the speaker says she's not being punched, she's doing the punching. Guntrip observes that

his patient has identified with, and therefore adopted the perspective of, a punitive mother, while retaining the aspect of the detested child she at bottom feels herself to be, a 'child within' who maintains a masochistic relationship with an internalized mother. He mentions several motives: identifying with the punitive mother provides relationship with her; it represents the struggle to preserve an ego; it confers a sense of power, even if only over her self. Guntrip uses Fairbairn's conceptual framework which is more revealing of this clinical case, but Freudian language will do, and we can describe the girl's situation without excessive distortion as her ego being maltreated by a punitive superego that perceives its object as a wretched, hateful thing. On the other side, the ego perceives the superego under the aspect of the punitive, though still consoling, mother. These 'parts of the self' perceive and treat each other as if they were alien objects. The patient thus appears to be divided in perspective, subjectivity, and agency, in the terms of partition when a 'real split' has occurred.[18]

How is it possible for such an intrapsychic sadomasochistically configured enactment to be wish-fulfilling? Briefly, one way is this: in its enactment the superego figure is manufacturing an experience of satisfaction (EOS) for the ego—and at the ego's behest, or in response to its needs. We might say, paradoxically on the face of it, that the ego activates the superego and the superego then enacts the consoling/punitive repertoires written into it. The ego, not knowing fully the superego's repertoire, may get more than it bargains for. (Recall the quasi-independence and opacity of the agencies.) Principally the ego seeks the companionship and security provided by the internalized figure, and identification with it as aggressor, even if that comes with a familiar dose of punishment. In re-establishing that relationship, of which we can observe only the punitive enactment, the person, qua ego, has her wishes fulfilled, in the manner of FW. For we can see how the ego figure may be engrossed or immersed in the EOS that is the superego's enactment. So here are circumstances analogous to Freud's basic model of FW. There is, as it were, a theatre of irrecusable evidence presented to the ego from which it cannot but draw wish-fulfilling conclusions: 'mother is present and I am her beating a child'. The case of the bedstead girl may not be dissimilar. The ritualistic performance enacted as an EOS by an obsessional part of herself is unconsciously understood by her child-self as the wish-fulfilling achievement of separating her parents—while at the same time pacifying and achieving an unstable identification with that obsessional (superego) part.

W(h)ither Wish-Fulfilment?

Although Freud was educated in the reductionist biophysics movement of the late nineteeenth century even his earliest psychoanalytic work ran counter to the scientific positivism and neurological psychiatry associated with that movement. In contrast to it Freud insisted on the causality of subjective motives, on the desiring, wishful, and affective aspects of mind, particularly of course on sexual motives and their development. The idea that desire or unfulfilled wish causes mental conflict and disorder is older than

Plato and proved over the centuries, but with Freud something entirely new appears: mental disorder is not just *caused* by desires in conflict; symptoms (and their kin) are compromise *fulfilments* of those desires—in the manner of FW. This is a strikingly novel idea, fundamental to the explanation of human behaviour. Freudian wish-fulfilment, 'the doctrine of the wish' (as it was early on and enthusiastically designated (Watson 1916)), is a deep statement about the ineluctability of human desire and of the will to satisfy it. When we cannot change the world we change ourselves: if the world will not accommodate our desire then we dream, phantasize, manufacture delusions, modify our bodies and perceptions as bogus evidence that desire *has* been satisfied.

Desire and this human way of dealing with it are in some neglect in contemporary human sciences. Psychiatry, neuroscience, and cognitive psychology have attenuated their conceptual resources, so that it is unclear whether questions about mental illness once thought to be intricated with human longing and despair, or willed delusion, can even be framed in them. They are now dominated by forms of reductive materialism and simplistic conceptions of motivation. And were it possible to provide *complete* descriptions of mental disorder, dream, phantasy (let alone artistic creation and religious devotion) in the languages of the brain sciences then their exclusion of almost all traces of the Intentional, common-sense-psychological, understanding of disorder, of which FW has been an important component, would be entirely vindicated. But, of course, it is not.

In classical psychoanalytic theory most mental disorders *were* considered largely as pathologies of desire and conflict, described in the Intentional idiom; but that has given way in contemporary theory to a dominant concern with pathologies of developmental deficits resulting from environmental, attachment, or interactional failure. Psychoanalysis now appears less concerned with the pathology of desire than of belief—or the various proposed non-Intentional but belief-like structures anteceding beliefs—internal working models (Bowlby), representations of interactions generalized (D. N. Stern), and so on. The attention to environmental deficiency and early impairment has, by and large, diminished interest in the desiderative, conflictual, and wish-fulfilling aspects of mind. The theoretical contrast is odd because trauma, privation, deprivation become mensurable mostly against a backdrop of need and desire; and the language of object or interpersonal relations is otiose without reference to mutual desiderative and affective relations. So of course in reality motivational notions do not disappear and FW reappears, as omnipotent phantasy, projective identification, manic defence, and so on. The scope Freud had given it has been circumscribed but wish-fulfilment is still shifting in the background, wearing different clothes, too much unremarked.[19]

References

Boag, S. (2011). 'Addressing mental plurality: justification, objections and logical requirements of strongly partitive accounts of mind'. In N. Mackay and A. Petocz (eds), *Realism and Psychology*. Boston, MA: Brill.

Bollas, C. (1987). *The Shadow of the Object*. New York: Columbia University Press.

Bowlby, J. (1998/1980). *Attachment and Loss*, Vol. 3: *Loss*. London: Pimlico.

Brakel, L. A. W. (2009). *Philosophy, Psychoanalysis, and the A-rational Mind*. Oxford: Oxford University Press.

Cavell, M. (1993). *The Psychoanalytic Mind*. Cambridge, MA: Harvard University Press.

Cavell, M. (2006). *Becoming a Subject*. Oxford: Clarendon Press.

Dodds, E. R. (1951). *The Greeks and the Irrational*. Berkeley, CA: University of California Press.

Eagle, M. (2011). *From Classical to Contemporary Psychoanalysis*. New York: Routledge.

Ellenberger, H. F. (1970). *The Discovery of the Unconscious*. New York: Basic Books.

Freud, S. (1894). 'The neuro-psychoses of defence'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 3 (pp. 43–61). London: Hogarth Press.

Freud, S. (1895). *Studies in Hysteria* (with Joseph Breuer). In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 2. London: Hogarth Press.

Freud, S. (1900). *The Interpretation of Dreams*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vols 4 and 5. London: Hogarth Press.

Freud, S. (1901a). *On Dreams*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 5. London: Hogarth Press.

Freud, S. (1901b). *The Psychopathology of Everyday Life*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 6. London: Hogarth Press.

Freud, S. (1908a). 'Hysterical phantasies and their relation to bisexuality'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 9 (pp. 155–166). London: Hogarth Press.

Freud, S. (1908b). 'Creative writers and daydreaming'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 9 (pp. 141–153). London: Hogarth Press.

Freud, S. (1909a). 'Some general remarks on hysterical attacks'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 9 (pp. 227–234). London: Hogarth Press.

Freud, S. (1909b). 'Notes upon a case of obsessional neurosis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 10 (pp. 153–249). London: Hogarth Press.

Freud, S. (1911). 'Formulations on the two principles of mental functioning'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 12 (pp. 213–226). London: Hogarth Press.

Freud, S. (1912). 'Contributions to a discussion on masturbation'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 12 (pp. 239–254). London: Hogarth Press.

Freud, S. (1913). 'The claims of psychoanalysis to scientific interest'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 13 (pp. 163–190). London: Hogarth Press.

Freud, S. (1915a). 'Instincts and their vicissitudes'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 109–140). London: Hogarth Press.

Freud, S. (1915b). 'The unconscious'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 159–204). London: Hogarth Press.

Freud, S. (1916–17). *Introductory Lectures on Psychoanalysis*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vols 15 and 16. London: Hogarth Press.

Freud, S. (1917). 'A metapsychological supplement to the theory of dreams'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 217–235). London: Hogarth Press.

Freud, S. (1922). 'Some neurotic mechanisms in jealousy, paranoia and homosexuality'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 18 (pp. 221–232). London: Hogarth Press.

Freud, S. (1923). *The Ego and the Id*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 19. London: Hogarth Press.

Freud, S. (1925). *An Autobiographical Study*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 20. London: Hogarth Press.

Freud, S. (1926). *Inhibitions, Symptoms and Anxiety*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 20. London: Hogarth Press.

Freud, S. (1927). 'On humour'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 21 (pp. 159–166). London: Hogarth Press.

Freud, S. (1933). *New Introductory Lectures on Psychoanalysis*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 22. London: Hogarth Press.

Freud, S. (1940). *An Outline of Psychoanalysis*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 23. London: Hogarth Press.

Freud, S. (1950). *Project for a Scientific Psychology*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 1. London: Hogarth Press.

Friston, K. (2010). 'The free-energy principle: a unified brain theory?'. *Nature Reviews Neuroscience 11*: 127–138. doi:10.1038/nrn2787.

Gardner, S. (1993). *Irrationality and the Philosophy of Psychoanalysis*. Cambridge: Cambridge University Press.

Guntrip, H. (1968). *Schizoid Phenomena, Object Relations and the Self*. London: Hogarth Press.

Hampshire, S. (1974). 'Disposition and memory'. In R. Wollheim (ed.), *Freud: A Collection of Critical Essays*. New York: Anchor Books.

Hobson, A., Hong, C. C.-H., and Friston, K. (2014). 'Virtual reality and consciousness inference in dreaming'. *Frontiers in Psychology 5*: 1133. doi:10.3389/fpsyg.2014.01133.

Hopkins, J. (1982). 'Introduction: philosophy and psychoanalysis'. In R. Wollheim and J. Hopkins (eds), *Philosophical Essays on Freud*. Cambridge: Cambridge University Press.

Hopkins, J. (1988). 'Epistemology and depth psychology'. In P. Clark and C. Wright (eds), *Mind, Psychoanalysis and Science*. Oxford: Blackwell.

Hopkins, J. (1995). 'Irrationality, interpretation and division'. In C. McDonald and G. McDonald (eds), *Philosophy of Psychology*. Oxford: Blackwell.

Hopkins, J. (2000). 'Psychoanalysis, metaphor and the concept of mind'. In M. Levine (ed.) *The Analytic Freud*. London/New York: Routledge.

Hopkins, J. (2012). 'Psychoanalysis, representation, and neuroscience: The Freudian unconscious and the Bayesian brain'. In D. Pfaff, A. Fotopoulu, and M. Conway (eds), *From the Couch to the Lab*. Oxford: Oxford University Press.

Hopkins, J. (2015). 'The significance of consilience: psychoanalysis, attachment, neuroscience, and evolution'. In S. Boag, L. Brakel, and V. Talvitie (eds), *Psychoanalysis and the Philosophy of Mind: Unconscious Mentality in the 21st Century*. London: Karnac.

Hopkins, J. (2016). 'Free energy and virtual reality in neuroscience and psychoanalysis: a complexity theory of dreaming and mental disorder'. *Frontiers in Psychology* 7: 922. doi: 10.3389/fpsyg.2016.00922.

James, W. (1950 [1890]). *The Principles of Psychology*. New York: Dover Books.

Lear, J. (1998). *Open Minded*. Cambridge, MA: Harvard University Press.

Lear, J. (2015). *Freud*, 2nd edn. New York: Routledge.

Maze, J. R. (1983). *The Meaning of Behaviour*. London: Allen and Unwin.

Mele, A. R. (2009). *Effective Intentions*. Oxford/New York: Oxford University Press.

Moore, M. (1984). *Law and Psychiatry*. Cambridge: Cambridge University Press.

Newberry, G. (2011). 'Drive theory reconsidered (again)'. In N. Mackay and A. Petocz (eds), *Realism and Psychology*. Leiden/Boston, MA: Brill.

Ogden, T. H. (1993 [1986]). *The Matrix of the Mind*. New Jersey: Jason Aronson.

O'Shaughnessy, B. (1980). *The Will: A Dual Aspect Theory*. Cambridge: Cambridge University Press.

Panksepp, J. and Biven, L. (2012). *The Archaeology of Mind*. New York: W. W. Norton.

Pataki, T. (2000). 'Freudian wishfulfilment and subintentional explanation'. In M. Levine (ed.), *The Analytic Freud*. London/New York: Routledge.

Pataki, T. (2003). 'Freud, object-relations, agency and the self'. In M. C. Chung and C. Feltham (eds), *Psychoanalytic Knowledge*. Houndsmill: Palgrave.

Pataki, T. (2014). *Wishfulfilment in Philosophy and Psychoanalysis: The Tyranny of Desire*. London: Routledge.

Pataki, T. (2015). 'Wish-fulfilment revisited'. In S. Boag, L. Brakel, and V. Talvitie (eds), *Psychoanalysis and the Philosophy of Mind: Psychoanalysis in the 21st Century*. London: Karnac.

Pears, D. (1984). *Motivated Irrationality*. Oxford: Clarendon Press.

Petocz, A. (1999). *Freud, Psychoanalysis and Symbolism*. Cambridge: Cambridge University Press.

Rorty, A. (1988). 'The deceptive self: liars, layers and lairs'. In B. P. McLaughlin and A. Rorty (eds), *Perspectives on Self-deception*. Berkeley, CA: University of California Press.

Rosenfeld, H. (1987). *Impasse and Interpretation*. London: Tavistock.

Sandler, J. (1976). 'Dreams, unconscious fantasies and "identity of perception"'. *International Review of Psychoanalysis 3*: 1.

Sandler, J. and Sandler, A. M. (1978). 'The development of object-relationships and affects'. *International Journal of Psychoanalysis 59*: 2.

Segal, H. (1991). *Dream, Phantasy and Art*. London: Tavistock Routledge.

Solms, M. (2015). *The Feeling Brain*. London: Karnac.

Solms, M. and Kaplan-Solms, K. (2000). *Clinical Studies in Neuro-Psychoanalysis*. London: Karnac.

Spitz, R. (1965). *The First Year of Life*. New York: International Universities Press.

Steiner, J. (1993). *Psychic Retreats: Pathological Organizations in Psychotic, Neurotic and Borderline Organizations*. London: Routledge.

Watson, J. B. (1916). 'The psychology of wishfulfilment'. *The Scientific Monthly* Nov. 1916.

Weiss, J. and Sampson, H. (1986). *The Psychoanalytic Process*. New York: International Universities Press.

Winnicott, D. W. (1965). *The Maturational Process and the Facilitating Environment*. London: Hogarth Press.

Wollheim, R (ed.) (1974). *Freud: A Collection of Critical Essays*. New York: Anchor Books.

Wollheim, R. (1979). 'Wish-fulfilment'. In R. Harrison (ed.), *Rational Action: Studies in Philosophy and Social Science*. Cambridge: Cambridge University Press.

Wollheim, R. (1984). *The Thread of Life*. Cambridge: Cambridge University Press.

Wollheim, R. (1991 [1971]). *Freud*. London: Fontana.

Wollheim, R. (1993). *The Mind and its Depths*. Harvard, MA: Harvard University Press.

Wollheim, R. and Hopkins, J. (eds) (1982). *Freud: A Collection of Critical Essays*, 2nd edn. New York: Anchor.

Notes:

(1.) Father Rageneau, *The Jesuit Relations and Allied Documents* (1897–9), quoted in Ellenberger (1970: 26).

(2.) There is a sense of wish-fulfilment independent of belief and satisfaction, as when a deceased person's wishes are fulfilled after his Will passes probate. We can leave this sense aside.

(3.) For a summary see Freud (1913: 186–7). The 'actual neuroses' were early exceptions to the wish-fulfilment formula and later Freud made accommodations for traumatic dreams and symptoms, but the general view held consistently despite major changes to other aspects of theory.

(4.) The common-sense-psychological (CP) classification of motives and acts used in this chapter is as follows. *Intentional action* is caused and rationalized by the conjunction of desire and instrumental belief; it may also involve executive states such as intending, deciding, and choosing; but the having of reasons is both causally and logically necessary. Various non-intentional modes of causal explanation are available to CP. O'Shaughnessy (1980) highlighted the notion of *subintentional action*. Such acts, including mental acts, are caused by desire without facilitation by instrumental beliefs. There are forms of *expression* that have archaeology but no teleology, as when one clenches fists in (out of) anguish, or laughs. Explanation along these lines belongs in CP but is neither intentional nor subintentional. As Richard Wollheim and Jim Hopkins first pointed out, much of Freud's work can be seen as an extension and deepening of this common-sense psychological framework. Useful, if not unproblematic, classification of such extensions may be found in Wollheim (1991: xixff; 1993).

(5.) The dreams interpreted in the temples of Asclepius in which a respected personage or god, 'reveals without symbolism what will or will not happen, or should or should not be done' (Dodds 1951: 107), or those like Alexander the Great's Satyr dream (Freud 1900: 99), have a wishful tincture but they are obviously prognostic and instructive not wish-fulfilments or *substitutive satisfactions* in the Freudian sense.

(6.) Freud's implicit notion that an uncontradicted idea is *ipso facto* believed has pedigree (W. James (1950: II, 288) cites Spinoza as the source) but it remains live in the neuroscience of our day: 'In the absence of sensory constraints, the vivid percepts, delusional beliefs and cognitive defects cease to be delusional or defective—because these attributes are only defined in relation to sensory evidence. However, in sleep, there is no sensory evidence and the only imperative is to adjudicate and select among unconstrained scenarios that can be entertained by the sleeping brain' (Hobson, Hong, and Friston 2014). See also Solms (2015: 138–9). For elaboration of the notion of engrossment, introduced in Wollheim (1979), see Pataki (2014: 35–42).

(7.) Petocz (1999; this volume) distinguishes between Freud's narrow theory of symbolism (FN) of fixed archaic relations and a broader theory (FB). I accept that distinction. My concern here is to distinguish both these sorts of symbolism from condensation and displacement, vicissitudes that Freud also conceived as representational.

(8.) Petocz (this volume) modifies her earlier view and no longer contends that a belief in the identity of symbol and symbolized is essential to the motivated mistaken identity that underlies the capacity to treat one thing as if it were another. I present the earlier view because it still seems to me a live option, one of several that may facilitate symbolic FW, as the discussion in the section 'Intention in Wish-Fulfilment' shows (see also Pataki (2014): 88–94). Although she does not explicitly address the issue of symbolic wish-fulfilment in this volume, the argument advanced for the compatibility of drive discharge theory with 'cognitive' or 'informational' approaches to drive activity suggests than an account of FW in terms of substitutional drive discharge, where the symbol is mistaken for the drive's primary object, and the drive is figured as a complex entity with cognitive capacities, will suffice. My argument for the necessity in FW of the occurrence of a suitable EOS or wish-fulfilling belief suggests that it will not. In any case, I cannot follow Petocz (1999: 220–32; this volume; Maze 1983; Newberry 2011) in the constructions placed on Freudian drive theory. Maze, and in earlier work Petocz, argued that drives are programmed with innately, though highly modifiable, consummatory activities, that they must have the capacity to seek out and perceive the conditions that terminate drive activity, that these cognitive states are components of drives: in short, drives are 'knowers'. In the present chapter the relevant psychological states (perceiving, believing, remembering) are described as 'crucial *evolved components* of drive-consummatory activity', not as components of the drive itself, but a somewhat 'thick' view of drives is retained and it leads Petocz to favourably consider the implausible proposition that the psyche or self is 'composed' (exhaustively constituted?) of instinctual drives. These views seem to me neither accurate renderings of Freud's account of drives—ambiguous as that is—nor sustainable. The fact that drives causally interact with cognitive states—or that perceptions and beliefs influence our desires and behaviour—does not entail either that cognitive states are components, mereological parts of drives, or that drives can sense, peer at, or know anything. It must be conceded however that the 'thick' view of drives has the appeal of potentially resolving some of the many conundrums of mental plurality to which Petocz draws attention. (See also Boag 2011). I propose a different approach to some of these conundrums in the sections 'Intention in Wish-Fulfilment' and 'Self-Solicitude'.

(9.) This is discussed further in the section 'Intention in Wish-Fulfilment'. For reasons such as these Sandler (1976; Sandler and Sandler 1978) postulated an 'understanding work' running parallel with but in opposite direction to the distortions of dream and symptom. Thus the manifest or surface expressions are 'decoded' in the Pcs. of the topographical theory. or unconscious ego of the structural theory, and the latent meaning is then 'understood'.

(10.) I am assuming (the contested view) that the explanation of the causal efficacy of conscious Intentional and phenomenal properties qua mental properties remains problematic for non-reductive physicalism, the metaphysical view implicit in the work

under discussion. Conscious experience 'fits' in this enlargement of the Helmholtz-Friston account, and is made to do work, but is not, as it seems to me, explained.

(11.) Those possibilities are explored in Pataki (2014, 2015).

(12.) I use 'motive' loosely to designate anything that can be a cause of action. 'All the categories', Freud wrote, 'which we employ to describe conscious mental acts, such as ideas, purposes, resolutions and so on, can be applied to them [unconscious mental acts]. Indeed, we are obliged to say of some of these latent states that the only respect in which they differ from conscious ones is precisely in the absence of consciousness.' (Freud 1915b: 168). Freud invokes unconscious motives to explain the wish-fulfilling phenomena of symptoms (e.g. 1916–17: ch. xix–xiii, esp. 298–9; 1909a: 231); parapraxes (1901b); phantasies (1908a: 159, 164; 1909a); play (1908b: 146; 1911: 39); and art (1908b); daydreams and masturbation phantasies (1912). His description of neurosis as consoling flight into illness is certainly suggestive of intentional strategy, and it could hardly have been plainer that this was his view of the secondary purpose in neurosis (1909a: 231–2).

(13.) On the counter-intuitiveness of intentional symptom formation see Eagle (2011: 70). The intuition that 'intention is the one concept that ought to be preserved free from any taint of the less-than-conscious' (Hampshire, 1974: 125) is widely shared but currently much disputed (Mele 2009). Unconscious agency receives experimental support in Weiss and Sampson (1986) and neuropsychoanalytic support in Kaplan-Solms and Solms (2000: esp. 108, 111). For problems with partitive conceptions of mind see Gardner (1993). Some notable non-intentional readings of FW within a broadly common-sense psychological framework: Wollheim (1984, 1991, 1993), Hopkins (1982, 1995), Moore (1984), Lear (1998, 2015), Gardner (1993), Cavell (1993, 2006), and Brakel (2009). The larger canvas is in Pataki (2014, 2015).

(14.) [Eds: Cf. Leite (this volume) for further exploration of this theme.]

(15.) Spitz (1965: 179) noted that internalization of maternal functions is essential for the infant's capacity for self-regulation. Winnicott (1965: 48): 'The infant develops means for doing without actual care. This is accomplished through the accumulation of memories of care, the projection of personal needs and the introjection of care details'. To my knowledge the implications for mental structure of such normal internalization of caring functions has not been much investigated, though see Bollas (1987); Kaplan-Solms and Solms (2000: 234). More familiar are the pathological dissociations associated with the internalization of frustrating or hated objects, or with false-self formation based on compliance or denial. See Winnicott (1965); Rosenfeld (1987); Steiner (1993); Bowlby (1980). It is possibly because the self-solicitous forms of FW are so frequently dominated by pathological developments that they often issue, paradoxically, in self-deceit and symptom.

(16.) I use 'self' and 'parts of self' promiscuously, counting on a measure of disambiguation by context.

(17.) Freud: 'We were justified, I think, in dividing the ego from the id, for there are certain considerations which necessitate the step. On the other hand the ego is identical with the id, and is merely a specially differentiated part of it. If we think of this part by itself in contradistinction to the whole, *or if a real split has occurred between the two*, the weakness of the ego becomes evident … The same is true of the relation between ego and superego. In many situations the two are merged; and as a rule we can only distinguish the one from the other when there is a tension or conflict between them'. (1926: 97; my italics). Amongst philosophers, to my knowledge only David Pears (1984), Rorty (1988), and Pataki (2003, 2014) have argued for a conception of mind whose parts satisfy conditions for independent agency. Within psychoanalysis, on the other hand, it is commonplace to hold that the mind is constituted by independent agencies. Freud, Fairbairn, M. Klein, Winnicott, Bion, Ogden, and S. Mitchell explicitly embrace partitive conceptions of mind. A roll call doesn't settle anything about the adequacy of partitive conceptualization but it can suggest that the clinical phenomena press strongly in its direction.

(18.) I am of course taking a highly Realistic line on the agencies. The defence with a more accommodating conceptual framework is set out in Pataki (2014: ch. 5). Casting the agencies involved as person-like has the happy consequence that the language of personal relations, the Intentional or common-sense-psychological idiom, can be enlisted in the description of intrapsychic relations.

(19.) This chapter has been greatly improved by the meticulous commentary of the editors of this volume, and by discussions with Jim Hopkins and Agnes Petocz. I am much indebted to them.

Tamas Pataki

School of Historical and Philosophical Studies, The University of Melbourne

Oxford Handbooks Online

Abstract and Keywords

The very idea of psychic integration presents puzzles in the case of unconscious belief, both for the analysand and for the theorist. In many cases, the unconsciously believed proposition is one that the analysand knows perfectly well to be false. What could it be to bring such a belief to consciousness? What could psychic integration come to in this sort of case? Put bluntly, the task facing the analysand is to consciously hold the belief even while placing it within a broader perspective in which it is recognized to be false. Implications are drawn concerning a number of large issues in epistemology and philosophy of mind: Moore's Paradox, the role of rationality in psychic unity and self-consciousness, the nature of the first-person standpoint in relation to one's own attitudes, transparency accounts of self-knowledge, and the role of endorsement in the constitution of the self.

Keywords: unconscious belief, psychic integration, Moore's Paradox, rationality, self-consciousness, first person, self-knowledge, transparency, self-constitution

Adam Leite

> 'we find that unconscious beliefs control our patients' lives and they are therefore important to discern'
>
> Ronald Britton, as reported by Paul Williams (1996: 81)

Introduction

THERE are as many forms of psychic integration as there are forms of psychic disorganization and fragmentation. It's one thing to bring together feelings that have been held apart, another to take back projections, and still another to be able to represent previously unrepresented mental states or to bring psychosomatic phenomena more fully into one's mental life. My focus in this chapter is on just one form of

integration: the integration of psychodynamically unconscious belief. This issue looms large in clinical practice (Britton 1998: 8). It also has rich and surprising philosophical implications.

A step towards psychic integration is a step towards a more stable and more inclusive psychic organization from a previous disorganized, fragmented, or incomplete state. Stable organizations are possible that are wildly dysfunctional and out of touch with reality (Steiner 1993). Insofar as increasing integration is understood as a move towards psychic health, it must take place in a way that involves increasingly apt response to reality. At the ideal, 'integration' is a success term. It requires both getting things right enough about oneself and the world and also bringing these matters into contact with the rest of one's psychic functioning. In cases at issue here it involves insight, experientially based self-understanding (Bell and Leite 2016), self-acceptance, and acceptance of reality. A key question, then, forms the topic of this paper: What would all of this look like when it comes to certain psychodynamically unconscious beliefs that are—as the analysand can correctly see—plainly false?

(p. 306) My discussion will assume both a psychodynamic conception of the mind and a clinical picture on which integrating psychodynamically unconscious materials can be a step towards psychic health. My goal is to show what can be learned philosophically from taking this picture seriously. The discussion will be organized into the six following sections: 'Unconscious Belief' introduces the key notion under discussion; 'Integrating Unconscious Belief: The Problem' highlights a fundamental puzzle concerning the integration of unconscious belief; and 'Moore's Paradox and the Integration of Unconscious Belief' offers an initial solution. 'Is there a Superior Alternative Description?' considers and rejects various ways of redescribing the phenomena so as to avoid the puzzle. 'What Integration might Look Like' aims to make such integration intelligible, and, finally, 'Philosophical Implications' highlights important implications for several large issues in philosophy: the role of rationality in psychic unity, the nature of the first-person standpoint on one's own attitudes, the role of transparency in self-knowledge, and the place of endorsement in the constitution of the self. The upshot is a distinctive orientation towards these issues, one that may be said to be 'anti-rationalist' from the perspective of a common but overly narrow conception of what human rationality comes to.

Unconscious Belief

I will take for granted that there are beliefs which are unconscious for dynamic reasons relating to deep wishes, fears, and defences. For example, an analysand might unconsciously believe that he is the most delightful and most beloved of the analyst's patients and that their relationship is special in a way that even the analyst's marriage can't match. Another might unconsciously believe that her analyst has no interest in her life and struggles. These unconscious beliefs play a hidden role in shaping the person's conscious thoughts, desires, emotional responses, and behaviour. Conscious reflection

and deliberation generally will not modify them or otherwise affect the role they play. For these reasons, among others, it can be important for the analysand's positive development that such beliefs become conscious, in a sense of that term (discussed later) that is closely connected with self-awareness and the ability to self-ascribe the belief from a first-person standpoint.

In describing these states as *beliefs*, we take it that they are content-bearing states of the person that play a functional and explanatory role similar enough to that of ordinary belief to warrant lumping the two together. In characterizing them as *unconscious*, however, we highlight an important difference.

To bring the difference into view, it's helpful to have the notion of the *subjective perspective* of a belief. When somebody believes some proposition p, p's being the case is part of the subjective perspective through which and against the background of which the person encounters the world. We might say that in *her* subjective world, p is the case. Occupying this subjective perspective will involve patterns of emotional response and expectation, dispositions towards reasoning and action, certain phenomenal dispositions, (p. 307) and a tendency to seek corroborative evidence and to experience the world in ways congruent with p's being true.[1]

In ordinary conscious belief, the person is capable of occupying the subjective perspective of the belief *as conscious subject*. The world will be presented to and in her conscious occurrent thought as including *that p is the case*, and this presentation will include a distinctive phenomenology, which we might term the force of reality. In addition, ordinary conscious belief on the part of mature human beings involves *self-consciousness*, in the sense that the person can simultaneously *be* the conscious subject of the belief and also self-ascribe the belief from that very position. She can, we might say, occupy the attitude as her own.

In qualifying a belief as *unconscious*, then, we are highlighting that though the person is in a state which plays at least a good part of the functional and explanatory role of belief, the person currently either lacks the ability to occupy the subjective perspective of the belief as conscious subject or else lacks the ability to self-ascribe the belief from that position.[2]

For example, imagine someone with the following complex pattern of responses. He finds himself feeling guilty when kindnesses or favours are done for him, as though he has taken something away from the other person. He feels that he is in danger of being caught taking something he doesn't deserve if he accepts positive offers from others, and he experiences his legitimate and unproblematic requests for help as manipulative manoeuvres robbing others of their riches. He fears that accolades that come his way are nothing but the gains of trickery and deceit. He feels a vague terror that if others knew the truth about him they would hate and cast him out, so that he must always hide the truth about himself. It might be that what unifies all of these reactions is this: he believes that he is nothing but a thief. But though he has all of these thoughts and feelings both in particular situations and as part of the ongoing background of his life, he still might not

consciously encounter the world as containing this truth: *I am a thief*. In such a case, he cannot consciously occupy the subjective perspective of the belief, and he cannot self-ascribe it from that position. His belief is functioning unconsciously.

(p. 308) In understanding him in this way we attribute to him a belief content that conflicts *in his own thinking* with thoughts to the effect that he has not stolen. This too is supported by the detailed texture and particular pattern of his responses. For instance, a range of very particular emotional reactions (such as specific forms of guilt connected in various ways with the idea of theft) are forced upon him, while others (notably such attitudes as pride and contentment) are systematically excluded or distorted. He likewise interprets particular incidents in ways that fit with the thought that he has stolen, while conflicting facts are explained away, reinterpreted, or ignored. When told that he did well at something, he will react in a variety of ways that express, maintain, or even reinforce feelings of guilt connected with thoughts about theft. He might disavow that *he* really did the thing (insinuating that he stole credit for it) or deny that he did it *well* (insinuating that he somehow stole in accomplishing what he did). He might find a way to treat the compliment itself as stolen, e.g. by interpreting himself as having somehow 'fished' for it. As a last resort he might respond by defending against feelings of guilt and then feel that he is engaged in a cover-up. What he cannot do is sit comfortably with the thought, 'I just did well at such-and-such, and this person has, quite reasonably, voluntarily complimented me for it'. What we see here is a standing orientation towards himself involving feelings of guilt and thoughts about being undeserving, taking things from others, and the like, all of which he routinely connects with ideas of theft.

We gain a measure of understanding if we attribute a mental state to him that dictates these particular forms of response, an attitude that distorts his experience of himself and the world in ways that fit—and indeed insist upon—the truth of this proposition. This parallels theories of psychological attribution that emphasize interpretive fit with the overall pattern exhibited in the person's thought, feeling, and behaviour (cf. Child 1996; Davidson 1970: 221–2; 1983; 1984; Freud 1917; Mölder 2010). The pay-off is a gain in intelligibility, or as Freud terms it, 'a gain in meaning' (Freud 1915: 167).

Such psychological attributions are neither indefeasible nor unrevisable. In the therapeutic setting the attribution is underwritten by the ongoing fine details of the analys and's associations, thoughts, feelings, and reactions in interaction with the analyst. It might be refined retrospectively in light of later developments and the patient's settled understanding. In other cases, the state itself might change and develop over time.[3]

Many unconscious beliefs are unconscious for motivational reasons: that is, psychodynamic factors involving wishes, fears, and defences. The attribution of a particular psychodynamic explanation in a particular clinical case is supported in part by the way in which things unfold as the analytic process proceeds—for instance, defensive processes and patterns of resistance, the way in which the belief comes into consciousness (p. 309) when a particular defence or sense of unbearable conflict is

gradually undone, and the context provided by the ongoing development of a broader understanding of the patient's mind.

There are various routes by which unconscious belief can be psychodynamically generated. In some cases a recognized truth may be repressed because it is unbearably painful. In this way, traumatic real-world experiences can give rise to unconscious beliefs.

In other cases, a false belief may be psychodynamically generated. For instance, a belief may arise as a wish-fulfilment or as a defence against some other unbearable or unacceptable experience or psychic content. Britton (1995) describes a psychotic patient who believed that if she did not see her mother, she would go blind. This was a 'defensive counterbelief' formed to head off fear about what might happen to her mother when absent. (The protection comes from the childish thought, 'She's not absent; rather, something is wrong with my eyes and so I cannot see her'.)

Defensive counterbeliefs—like wish-fulfilling beliefs—can be conscious or can be kept unconscious because they are painful, threatening, conflict with central parts of one's conception of oneself or loved ones, and the like. In other cases, the belief is unconscious not so much to ward it off as to maintain it: the belief functions unconsciously as a way of safeguarding it in the face of acknowledged contradictory facts which are themselves felt (perhaps unconsciously) to be threatening or unacceptable.[4] Sometimes both factors are in play.

Importantly, psychodynamically unconscious beliefs (such as unconscious defensive counterbeliefs) need not be part of any larger rational structure. These beliefs exert rationalizing pressure insofar as they lead the person both to interpret situations in ways that conform with them and to fabricate justifications for convictions arising from them. But such considerations are not what holds them in place. Rather, they are held in place by motivational forces, mechanisms of defence, and other psychodynamic factors.

Two other important psychoanalytic notions—unconscious phantasy and implicit procedural schemata—stand in complicated relations to the concept of unconscious belief.

First, unconscious phantasy.[5] Different psychoanalytic traditions diverge in their interpretation of this protean concept (Bohleber et al. 2015). Frequently likened to daydreaming, phantasy is widely understood as 'an imagined scenario or storylike narrative' that represents the subject's interactions with others in ways 'shaped by motivational states and defensive operations' (Auchincloss and Samberg 2012a: 85). On many views, unconscious phantasy can lead to or produce both conscious and unconscious beliefs. There is disagreement, however, about several matters relevant to unconscious phantasy's relation to belief, including its origins, its relation to sources of motivation, (p. 310) and the nature of its content (propositional, conceptual, imagistic, or affective/bodily sensation).

Many prominent conceptions sharply distinguish the two. In a canonical Kleinian formulation, Susan Isaacs writes:

> phantasies are the primary content of unconscious mental processes … this 'mental expression' of instinct *is* unconscious phantasy. Phantasy is (in the first instance) the mental corollary, the psychic representative, of instinct. There is no impulse, no instinctual urge or response which is not experienced as unconscious phantasy.

> (Isaacs 1948: 91)

For Isaacs, unconscious phantasy is the first, 'low-level', mental manifestation of a drive or motivational urge ('instinct', in her terminology, following Strachey's translation of Freud's '*Trieb*' (drive, urge)). It likewise constitutes the mental content of defensive processes countering these urges. It has a non-propositional, affectively laden representational content (Gardner 1993: sect. 6.5). It thus differs from belief in several respects, including the fact that (unlike wish-fulfilling belief) it is not a *product* of the urge, but rather its mental expression.

Britton, too, sharply distinguishes unconscious belief from unconscious phantasy, holding belief to be what bestows the status of reality upon the phantasy's content for the subject (1995: 19–20; 1998: 9).

Writers in other traditions bring the two closer together. Erreich, for instance, suggests that unconscious phantasy should be 'defined as a more or less unconscious belief statement' (2003: 569) that results from defensive processes of compromise formation operating upon three components: wish, veridical perception, and naive cognition. However, Erreich also characterizes unconscious phantasy as 'a representational structure for mental content with a motivational component' (2003: 569) and views it as 'a vehicle for the mental representation of … affects, wishes, defenses' (2003: 545). So understood, it is doubtful that an unconscious phantasy simply *is* a belief, despite Erreich's explicit definition; it might perhaps be understood as a complex structure involving motivational elements along with unconscious beliefs as constituents, components, or 'concomitants' (2003: 567), or it might be understood as some sort of motivationally and affectively laden hybrid.

Some prominent views do not take a clear stand on the issue. In an influential contribution to the American ego psychology tradition, Arlow (2008/1969) characterizes unconscious phantasy as 'unconscious daydreaming' that is an 'ever-present accompaniment of conscious experience' (2008/1969: 42). If we take the identification with daydreaming literally, then unconscious phantasy is not a form of belief, even if it is 'composed of elements with fixed verbal concepts' (2008/1969: 23). Arlow also writes, however, that 'what is consciously apperceived and experienced is the result of the interaction between the data of experience and unconscious fantasying' (2008/1969: 42), and he likens unconscious phantasy activity to a source of input analogous to that (p. 311) provided by the senses (2008/1969: 43). He additionally characterizes unconscious

daydreaming as a mental activity that 'supplies the mental set in which the data of perception are organized, judged, and interpreted' (2008/1969: 43). These formulations leave it unclear whether and to what extent unconscious phantasies are either belief-like or involve a belief component.

A second explanatory construct can be more clearly distinguished from the concept of unconscious belief. Some psychoanalytic theorists have recently emphasized nonlinguistic procedural schemas, models, or routines for interpersonal relating—psychological structures which are stored in an implicit (non-declarative) memory system and explain patterns of behaviour, thought, and emotional response by generating particular dispositions, motivations, and mental contents (including beliefs) in particular contexts (for instance, Beebe and Lachmann 1994, Fonagy 1999, Lyons-Ruth 1999, Stern et al. 1998; for discussion, Eagle 2013). Such schemata may be thought to be underlying psychological realizers of some unconscious beliefs. Whether that is so depends on larger debates in the philosophy of mind.

In any case, however, the two concepts are distinct and are not extensionally equivalent. First, many implicit procedural schemata—such as govern distance–standing as well as more complex interpersonal behaviour—are not psychodynamically unconscious. Rather, they remain difficult to articulate because stored in implicit (non-declarative) procedural memory (Fonagy 1999: 216–17; Stern et al. 1998: 905–6). They consequently have different behavioural manifestations. The patterns of resistance, defensive response, and shifts over time that will lead an analyst to attribute a psychodynamically unconscious belief are present only in certain cases. Moreover, as noted earlier, some unconscious beliefs are generated by processes of psychic defence. Some of these can have a very different sort of clinical manifestation and profile, and can play a different explanatory role, from that of implicit procedural schemata. The two categories are thus disjoint. However, none of this is to deny that procedural schemata for interpersonal relating can participate in psychodynamic processes and structures: for instance, they can sometimes generate psychodynamically unconscious phenomena, remain unacknowledged for psychodynamic reasons, be shaped by psychodynamic processes, and even involve characteristic interpersonally enacted defences (Lacewing, this volume).

Integrating Unconscious Belief: The Problem

Much clinical theorizing emphasizes that it can be crucially important for a patient's progress that unconscious beliefs become conscious. As Britton puts it, 'until a patient discovers that he really believes an idea its correspondence, or lack of it, with external reality is irrelevant. His belief remains suspended, unconscious, unmodified, and unverified, in parallel with a contrary belief' (1995: 21).

(p. 312) What matters here is not simply awareness that one has a particular belief. An analysand might be said to 'discover that he really believes an idea' in the sense that he learns to recognize the signs of his unconscious belief as they show up in his thought, speech, and behaviour. On this basis he can say, 'I believe that so and so'. However, such

a person remains no closer to his belief *as conscious subject* than he was before (Moran 2001; Finkelstein, this volume). This might be a stage in a positive development nonetheless, perhaps even a step towards bringing the belief into consciousness. The person might benefit from learning to predict and manage the effects of the belief even if it never becomes fully conscious. Still, this is fundamentally different from being able to consciously occupy the subjective standpoint of the belief and self-attribute the belief from precisely that position. It is only when one has attained that latter ability that one can self-reflectively recognize the belief in operation 'from the inside' and bring it into contact with one's ongoing thinking about the matters with which it is concerned.

Bringing an unconscious belief into consciousness does not merely involve coming to occupy this position in relation to that single belief. Rather, it involves bringing the particular belief into the larger system of one's conscious thoughts, feelings, beliefs, and desires. It thus both requires and constitutes a process of psychic integration—a process of bringing together into a single perspective what has been held apart or cut off, for instance in a process of psychic defence. This necessarily involves an increase in the range of positions—thoughts, feelings, and beliefs—that one can simultaneously self-consciously occupy as a conscious subject.

To approach what might be involved here, it can be helpful to consider a parallel schematic example. In the defence known as 'splitting', conflicting emotional responses—along with correlative contradictory representations of self and other—are kept distinct (Auchincloss and Samberg 2012b; Kernberg 1966: 238). The defence operates to evade the anxiety introduced by the felt conflict between the emotions, and it can be exacerbated by the analysand's sense of threat arising from intense feelings of hatred (Klein 1946). If an analysand has split her feelings of love and hatred, alternately experiencing herself as in loving communion with an ideal object and in fierce battle with a denigrated object, it will be crucial for her positive development that she come to experience her hatred, understanding it as such, *while also remaining in contact with her loving feelings*. That is the work of integration in this sort of case. The difficult task facing the analysand is to hold in mind the feelings of both love and hate, understood precisely as love and hate, and to experience their conflict as such, thereby uniting them into a single self-conscious perspective. This amounts to bringing her loving self and her hating self together into one.[6]

Integration in this sense does not necessarily entail rational control. Someone who manages to integrate her love and hatred may not be able to get rid of either by reflecting on the reasons for and against them. She may not be able to modulate the one by 'mingling' (p. 313) it with the other (and that may not be appropriate in any case). Still, her new ability to hold both in mind will enable new forms of thought, feeling, and reaction, including the possibility of experiencing her behaviour and motivations in new and emotionally richer ways, asking new questions about herself and others, and considering paths of action that formerly were not live options. This requires the ability to

recognize the conflict without becoming overwhelmed by anxiety, so that she can live within and from that position of conflict as a self-conscious subject.

Turn now to the case of dynamically unconscious belief. In many cases the belief is one that the analysand would perfectly well recognize as *false*.[7] For instance, consider again the belief *I am a thief*. Despite the reactions manifesting this unconscious belief, the person might know perfectly well that he is no thief and has earned his salary and received loving gifts and assistance from family and friends. Here, concretely, is what this might look like. If we ask him, 'Have you earned your salary?', he replies, fully sincerely, 'Yes, of course. I work very hard.' And yet he might feel an inarticulate sense of guilt and then begin telling a story about someone else who embezzled. If he is given a gift out of love, he might recognize the gesture and genuinely feel grateful. And yet he might also find himself feeling that he has gotten away with something. These feelings and thoughts do not do away with his knowledge any more than everyday obsessive worries do away with one's knowledge that one turned off the stove. In his self-conscious thinking and reasoning, he retains a firm grip on the truth.[8]

Such cases present a particular challenge within the lived experience of the analysand. The analysand must attempt to take up, as a conscious belief, something he knows to be plainly false. There is a difficulty here which is not merely a matter of defence against anxiety or psychic pain. But these cases also present puzzles for us as theorists. What would it *be* to integrate a plainly false belief? How can we make sense of the idea of such a belief's entering into the person's self-conscious subjective perspective?

First, integration does not entail elimination. It would be lovely, perhaps, if as soon as the troublesome belief began to come into consciousness it automatically ceased to exist. However, because of the belief's roots in wish, phantasy, and defence, this is not how things usually go. These motivational forces require a different sort of treatment. For the same reason, integration does not guarantee modifiability through deliberation. It is a mistake to suppose that as the belief enters self-conscious mental life, it must become directly and fully accessible to control, modification, or revision through conscious reflection (contra Parrott 2015; cf. Moran 2001).

At the same time, however, integration can't require radical loss of one's grip on reality. It *can't* be that a belief comes into consciousness only insofar as the person now self-consciously believes, without qualification, something that she previously consciously (p. 314) regarded as false. Consider, again, Britton's patient who believed that if she did not see her mother, she would go blind (1995: 22). Such beliefs exert a rationalizing force of their own, leading to the construction of a structure of delusional reasons and evidence. To accept such a structure without qualification, and to feel and act accordingly, wouldn't be a route to integration, but to psychosis—a severe break from reality (as in the case of Britton's patient). Likewise, though confusion is sometimes an important aspect of the process of working through, it would not be a positive development in this sort of case if the person's ongoing conscious state were to become one of genuine uncertainty or confusion about what is true.

For these reasons, to understand what it would be to integrate such a belief we need to make room for an initially surprising idea: that a person could *simultaneously* self-consciously believe something *and* recognize that the belief is false. That is, we have to allow for the possibility of at least a period in which the person recognizes that she believes that she has stolen everything good that has come her way—and occupies the subjective position of that belief, viewing the world in that way—even while also perfectly well recognizing that this belief is false because there are a great many good things in her life that she did not steal: some were earned through hard work, some were freely given, and sometimes she was simply lucky. This state may not be the ultimate goal—an issue to which I will return—but we have to allow for it as a possible and sometimes essential step along the way.

Moore's Paradox and the Integration of Unconscious Belief

What I have been describing is a psychological position that might be summarized in an utterance or affirmative judgement of the so-called 'commissive' version of Moore's paradox: namely, 'I believe that p, but p is false' (Moore 1993). Many contemporary philosophers take such an utterance or affirmative judgement to be paradoxical. While it can be true that someone believes something that is false, it can seem incoherent to self-ascribe a belief from the first-person standpoint even while judging or declaring that the believed proposition is false. This incoherence has seemed to many philosophers to go beyond mere irrationality; it is, we're told, a fundamental and defining feature of the concept of belief. As Wittgenstein writes, 'One can mistrust one's own senses, but not one's own belief' (2001 [1953]: 190), and 'If there were a verb meaning "to believe falsely", it would not have any significant first person, present indicative' (2001 [1953]: 162). Many recent philosophers (e.g. Coliva 2015; Heal 1994; Moran 2001) have followed Wittgenstein in holding that we can make no sense of the idea of someone's consciously occupying the subjective standpoint of a belief, self-attributing it from that position, and also acknowledging its falsity. And yet this is precisely what integration of dynamically (p. 315) unconscious beliefs often seems to require. The psychoanalytic case thus requires us to accept as actual precisely what this tradition deems unintelligible.[9]

To make sense of this, it's helpful to think more about what is involved in belief.

Consider the difference that is made when a belief functions in the way characteristic of conscious belief. Conscious belief that p constitutively involves that p's being the case is part of the subjective perspective through which and against the background of which the person consciously encounters the world. As noted earlier, the world will be presented to and in her occurrent thought as including *that p is the case*, and this presentation will include a distinctive phenomenology: a sense or feeling of reality that might be captured by saying that it is part of her subjective world that things are this way.[10] Her experience will likewise be shaped in various ways that fit with the truth of p, and she will be disposed to have congruent emotional and motivational responses. Central here too are

other differences in the person's dispositions. If she consciously believes that p, she is disposed to judge that p when she considers relevant questions. P will be presented in her conscious directed thinking as a premise for her reasoning about what to do and what is the case, and she will be disposed to so deploy it. She will be disposed to assert p when appropriate and to act on the basis of p.

All of this can be true even if a person also recognizes that her belief is false. Crucially, many of these dispositions are defeasible, and they can be defeated in various ways and to varying extents in different cases. In particular, someone might consciously believe that p, seeing the world in just the ways involved in consciously having this belief, and yet not deploy p as a premise in conscious reasoning, act on its basis, or assert it flat out. She might be disposed to judge p to be true and yet at the moment when she considers the question, judge it to be false. She might strive not to act in ways that are predicated on taking p to be true. Some of this may involve resisting these dispositions through an act of will, as when one consciously resists the strong temptation to act as if p were true. In other cases it may simply be a matter of what the person does in the course of her directed conscious thinking: she might simply think, 'But p is false', and that might be enough. These dispositions can also be defeated through changes in the person's underlying psychological functioning, so that the temptations and tendencies arising from these dispositions do not even show up in her conscious thinking.

(p. 316) Suppose, then, that the analysand recognizes both that there is in fact no evidence in favour of the thought that she is a thief and that all sorts of considerations support the contention that she isn't one. Suppose that she also understands that her belief *I am a thief* is itself a product of dynamic processes relating to deep wishes, fears, and defences, and suppose that she has some experientially based understanding of these wishes and fears and her defensive responses to them, so that all of this is functioning to some extent consciously as well. In such a case, the belief will be intelligible to her, no matter how wacky it is, in a way that goes beyond merely intellectual understanding, insofar as the belief links with emotions and motivations she has experienced 'from the inside'. In light of all this, she might experience herself and her world in all of the ways involved in believing that she is a thief, and yet she might attempt not to reason from *I am a thief* in her conscious thinking about what to believe, think, and do. She might not assert 'I am a thief', full stop. She might attempt to recognize when her emotional responses are born of this false belief, and she might attempt to treat them accordingly; likewise, with motivations that are given life by the belief. All of these further responses, along with the subjective perspective involved in believing that p, thus form a total complex perspective in which she self-consciously believes that she is a thief while not endorsing that belief and indeed while judging it to be false. Even as she sees the world through the lens of this belief, as it were, she also recognizes it as her lens, recognizes the ways in which it is inapt, and resists, suppresses, or otherwise does not engage in some of the patterns of response that would otherwise appropriately be involved.

Different psychological positions might meet this description, as I discuss later. For now, the point is this: we can in fact make sense of someone aptly capturing the reality of her overall psychological situation by simultaneously self-attributing a belief from the subject position and rejecting that belief as false. The key is to see that even as one occupies the subjective perspective of the belief as conscious subject, one's *total self-conscious perspective* at that moment might include a wider purview. While this is a rationally incoherent position, situations of this sort are familiar in ordinary life as well, particularly in cases in which one unexpectedly discovers that a view to which one is deeply attached is in error (Leite 2016).[11]

Any view of belief which emphasizes its connection with characteristic patterns of activity, thought, motivation, and the like, would lead us to expect exactly the sort of phenomena that I am pointing to. At least, it would do so unless one also holds the view (identified most prominently with Donald Davidson) that belief ascription requires that the person come out as *rational*, since I am proposing belief ascriptions that involve (p. 317) contradictory beliefs within the same total view. However, if we think of belief ascription as part of the attempt to *make sense of the person* or to *render her intelligible* —where we separate that notion from specifically *rational* intelligibility—then often precisely such a belief ascription will be the best fit. The distinctive form of intelligibility provided by person-level folk-psychological explanation is not primarily and narrowly rational intelligibility (Hursthouse 1991); a man's rubbing his face in his wife's sweater is made intelligible when we note that he is dealing with her recent death, and this intelligibility is not fully captured when we attempt to understand this case in terms of a 'rationalizing explanation' in the standard narrow sense. It rather has to do with considerations of meaningfulness and significance for the person: with the emotional resonances things have, the ideas with which they are associated, the motivations that are in play, and the like. Once we allow that these richer forms of intelligibility play a role in the attribution of attitudes such as belief, the possibility opens up of attributing much more complex, conflicted belief states to the person.

To sum up: I have been arguing that the formerly dynamically unconscious belief is integrated by being incorporated into a total conscious perspective that also contains its contradictory. The total perspective is thus rationally fractured. But the fracture is contained, in three senses. First, it is *limited*, in that it does not metastasize into a larger split that would introduce a division into the person's capacity for thought (as in compartmentalization or dissociation). Second, the fracture is contained insofar as the person is able to *hold the whole situation in mind*—both sides of the fracture, the fracture itself, and their relations, and so can experience the whole situation as an object of thought and as less threatening. Third, the various aspects of the person's overall perspective—including the formerly unconscious belief—are put in their places: they are given a distinctive organization in relation to each other. As a result, the person now functions differently. The person's overall response—given the contradiction—is as rational as it can be, insofar as the person aims not to act or reason on the basis of the false belief and will attempt to ameliorate its effects on her emotions and motivations.

This is the total complex situation that might be summed up and expressed by the utterance or thought, 'I believe that p, but p is false'.

It should be emphasized that this is *not* the mechanism that philosophers sometimes call 'fragmentation' (Lewis 1982). In fragmentation, two (or more) systems of belief are not allowed to come into rational contact with each other. But here, the formerly unconscious belief *is* brought into rational contact with the rest of the person's conscious thought even while being experienced as her belief, and it is rejected as false. That is precisely how it is integrated into her total subjective perspective: as an aspect of that perspective that is false. This integration is assisted by the way in which the belief is rendered intelligible through an experientially based, emotionally rich understanding of the dynamic sources of the belief—an understanding that arises through conscious experience of the fears and wishes that hold it in place, so that they too, as well as their relation to the belief, all become part of the single subjective perspective. To come to understand—on the basis of a rich conscious experience of the relevant emotions—that one's belief is, for instance, an infantile defensive reaction to infantile fears is (p. 318) simultaneously to render it intelligible, to accept it as what it is, and to put it in its place—even if it continues to be an ineliminable part of one's total perspective.

Is There a Superior Alternative Description?

I have been arguing that we have to allow for the possibility of self-consciously believing something that one recognizes to be false, if we are to make sense of the very idea of integrating the mind by bringing unconscious beliefs into consciousness. One possible response is to feel that my argument just can't be right: since (it is said) we can make no sense of self-consciously believing something one knows to be false, we must have been wrong to take this as a matter of belief to begin with.

In a well-known paper, the philosopher Tamar Gendler (2008) urges the explanatory importance of what she calls 'aliefs': automatic, associative, arational clusters of representational content, affect, and associated behavioural routines. It might be suggested that what I have been talking about are really *belief-discordant aliefs* that are brought into consciousness, not fully fledged beliefs that would generate a Moore-paradoxical position of the sort I have been describing. It might similarly be said that in many of these cases it would be more natural to express the attitude using *affective* rather than *cognitive* language, saying something like 'I feel like I've stolen everything good in my life, but of course I know that's not so'.[12]

In some cases, either of these forms of description might be apt. In particular, it might be the case that many of what psychoanalysts call 'unconscious beliefs' consist in associative clusters of affective and behavioural dispositions. From a clinical perspective, however, the important point is this. Even if they were to begin life as 'aliefs', dynamically unconscious beliefs can undergo a *developmental trajectory* in which they acquire determinate propositional content and are brought into relation with the rest of the person's conscious view of the world in ways that are emotionally and motivationally

transformative. It is precisely this work that is amongst the goals of the analytic therapy. But the result might well be a conscious attitude that conflicts with the person's other conscious beliefs.

Likewise, casting these phenomena in broadly affective terms is fine, so long as one does not lose track of the crucial fact that the attitudes in question *conflict with* one's settled view. So long as one keeps this point clearly in mind, the terminology does not matter. Where the terminology can become important is in the clinical setting, since talk of what one 'feels' can be used to distance oneself from acknowledging how one actually takes things to be. Gardner's characterization of defence is apt here: 'an operation on mental content that represents the cause of anxiety in such a way as to reduce or (p. 319) eliminate anxiety' (1993: 145). Talk of what one 'feels' can function defensively to avoid the experience of psychic conflict by disguising unwanted or unacceptable aspects of one's view of what is the case.

It would likewise be a mistake to think that what I have been describing is best understood as a situation in which it merely *seems* or *appears* to the person as if things are a certain way. For her to deny that she holds the belief and to go straight to 'it merely seems or appears to me as if I have stolen everything good', would be for her to deny something crucial. When the belief was functioning unconsciously, this *was* how things were for her, though she didn't recognize it. And she is attached to this way of viewing her world, insofar as doing so is fuelled by and satisfies particular motivations and defensive responses to them. Believing that things are this way thus does important psychic work for her. None of that necessarily changes as these matters come into consciousness. Thus for her to say that this is not belief (but rather a mere 'seeming' or 'appearance') would be for her to evade these truths about herself in a way that enables the processes that maintain the belief to continue unmodified and unexamined. For the same reason, to characterize her state as mere 'seeming' or 'appearance' would miss— and render unintelligible—the clinical fact, emphasized by Britton, that once a dynamically unconscious belief becomes conscious, a process of *mourning* is often required before the belief can be given up (1995: 22).

Similar points apply to the thoughts that the person merely has an 'inclination' to believe —but doesn't actually hold the belief—or that only 'a part' of her believes. To go straight to such self-characterizations might very well be a motivated attempt to evade the difficult conflicts presented by the states in question or to avoid approaching the motivations that hold the belief in place. This would enable one to maintain the conflicting view of things at a deeper level even while consciously avowing it in a mischaracterized form. Such a self-characterization would short-circuit the difficulties involved in integrating the false unconscious belief. Psychoanalytic-sounding ideas, too, can be used for defensive ends.[13]

Moore-paradoxical belief attributions also cannot be avoided by emphasizing implicit, non-propositional procedural schemata for interpersonal relating. All psychoanalytic writers in this vein grant that there is a place clinically for the analysand's conscious

articulation of implicit models for relating.[14] This cannot be accomplished except by the analysand's acknowledging from a first-person position how he takes things to be in his interaction with the analyst (e.g. 'I can't reach out to you for comfort, because you **(p. 320)** disapprove of comfort-seeking').[15] Since what is being articulated is how things *are* from the patient's subjective viewpoint, this amounts to the conscious expression of belief. And there is no way for the patient to fully own—from the inside, as it were—the full pattern of inhibition, motivation, emotional response, silence, etc. that is involved here, unless he recognizes this as his perspective. But in some cases this will involve a proposition which he regards as false.

An additional point is relevant here. Theorists who emphasize implicit procedural schemata do not hold that this theoretical construct can entirely replace the clinical role played by the concept of unconscious belief and its articulation in the process of working through, since these schemata can generate distinct psychodynamically unconscious beliefs. Fonagy, for instance, writes that 'Consciousness of the beliefs generated by such implicit memory models is crucial if the patient is to acquire the power to inhibit or modify them through the creation of a second-order representation of their inner experience . . .' (Fonagy 1999: 219). Here again, a puzzle about the integration of false unconscious belief will arise.[16]

Someone who thinks that the self-conscious self-attribution of false belief is just *impossible* might instead suggest that as the belief comes into consciousness, the analysand oscillates between incompatible, internally coherent viewpoints in order to find a resting place in one and give up the other. But this picture is not adequate either. First, there is often no larger viewpoint on the world available to the person within which the formerly unconscious belief could be made to look rational. Because the belief is sustained not by rational relations but rather by a tissue of phantasy, wish-fulfilment, and defensive processes, to give it a place in a larger rationally coherent view would require the wholescale creation of a massively delusive view of the world. Second, this picture is overly optimistic about the extent to which it is possible to rid oneself of such beliefs. Often they remain, and they need to be given a place within one's larger subjective perspective if one is to have hope of putting together a stable life. The picture of 'oscillation' cannot allow for this. Finally, to think that this is how things should *always* go is to maintain a fantasy of being able to get rid of psychic materials one does not want. It would be better in many cases to aspire to fully occupy one's subjective perspective with acceptance and understanding, where that very perspective includes a well-developed, emotionally rich, and motivationally effective conception of its own weaknesses and deficiencies.

I think that this conclusion is inevitable for any psychodynamic therapeutic approach that takes the patient's total first-person, subjective perspective seriously as a realm for exploration, articulation, and productive elaboration through speech and avowal.

(p. 321) What Integration Might Look Like

Various positive endpoints are possible. With work, the person may give up the belief and correctly understand herself as subject to an unshakeable emotionally resonant 'appearance' or 'seeming' arising from understood psychodynamic forces. Through underlying psychological shifts she might lose even this 'seeming' or 'appearance'. In other cases, the underlying motivations might remain but gain their satisfaction not in the person's cognitive life, not in the realm of seemings and belief, but rather in the realm of treasured conscious fantasy and play. However, in many cases the endpoint is one in which despite what the person might want, it would be untrue for her to deny that she believes that she is a thief. In these cases the false belief is consciously maintained but given its place within a broader perspective that acknowledges its falsehood, renders it intelligible, and ameliorates its effects.

Even here there can be variation along multiple dimensions. One such dimension is the precise *way* in which the person now occupies the subjective standpoint of the formerly unconscious belief.

To see this, consider that the ability to use 'I' in expressive self-ascription of the attitude doesn't yet guarantee incorporation of this belief into the right sort of larger self-conscious perspective. After all, someone could, at a given moment, consciously occupy the subjective position of the belief that he is a thief and say, in a melancholic, chastened tone, 'I'm an awful person; I deceive people in order to get their love. I'm nothing but a thief.' He might continue, in response to a therapist's comment, 'You believe that you are a thief', 'Yes, I do. I steal from everyone.' This person has not yet integrated his formerly unconscious false belief, because he is now *in the grip of it*. It isn't yet properly incorporated into the right sort of broader perspective. (This position might nonetheless be an important stage in a developmental process, and someone engaged in 'working through' may slip in and out of it.)

Someone might instead be able to self-attribute the belief from the subject position, consciously having the thoughts, feelings, and motivations that go along with it, and yet battle against it in light of its falsehood. He might fight and struggle to resist some of the dispositions it involves, for instance by effortfully correcting false interpretations it generates in particular interactions. He might be able to place all of this in the context of a broader understanding of the psychodynamic forces at issue and an ability to recognize them as they appear and function in situ. In these ways, he *manages* the false belief.

This situation is a step towards psychic integration, but precisely in its aspect of intrapsychic struggle, it threatens to break apart. To put it metaphorically, the belief that he is a thief is pressing to bust loose, and if it breaks free he might oscillate back towards a position in which he is in the grip of the belief. The person's overall mental state is thus *organized* at the moment, but it lacks a measure of *stability*. Like a peace imposed by armed garrison, it is subject to internal pressures that threaten to undo it.

A notable feature of this last case is that there is a sense in which the person restrains himself from fully occupying the subjective position of the belief, precisely because of the threat it is felt to pose. For instance, he might not be able to give voice to this belief by saying 'I am a thief' without sliding back towards being in its grip. While he can wholeheartedly self-ascribe the belief, he might feel enormous anxiety at thinking the first-order thought, 'I am a thief'. That belief's subjective perspective thus hasn't yet been stably and fully incorporated into a larger, unified perspective in which he simultaneously occupies the standpoint of the belief even while recognizing its falsity.

How could he give some form of first-order expression to the belief even while maintaining his hold on its falsehood? Imagine that while he recognizes that his belief is false, he is no longer engaged in intra-psychic struggle against it. Instead, he is able to say, with warm, ironic good humour and a twinkle in his eye, 'Of course I am a thief. After all, I'm doing it right now! I've gotten you to pay undivided attention to me for a full hour, and I've given you no love or attention in return.' Here, tone is everything. We have to imagine that the person is both giving voice to his belief that he is a thief and also *in the fact that he is speaking with ironic good humour* giving voice to what he would otherwise express by saying 'Of course, I'm not a thief'. He thus gives voice to both sides of the conflict while putting each in its proper place. And this single utterance doesn't just express the conflict; it expresses his total position regarding both the belief and the conflict it is involved in. For instance, his utterance also expresses, with good-humoured irony, the rationalizing pressure to seek confirming evidence for the belief that he is a thief ('See, I'm doing it right now!'), and the gentle, self-accepting, good humour with which all of this is said expresses his recognition that this belief (with all that it involves) is an immature solution to an immature problem. What we see here is thus an expression of the full, stable incorporation of the false belief into a single, unified self-conscious perspective.[17]

This person can give voice to both sides of the conflict without any oscillation in his total position—that is, without any change in his total mental state, including its associated dispositions and patterns of thought, feeling, and reaction. The utterance exhibits an apt, settled pattern of relations between both sides of the conflict and also an apt, settled orientation towards the conflict itself. The person thus has the ability to speak with one voice—a voice that expresses the totality of his complex overall perspective all at once. Indeed, it is this ability that is a key reason for describing this as *one* contemporaneous complex subjective perspective comprising both the subjective position of the belief and recognition of its falsehood. What underwrites such an ability are the functional relations amongst the person's beliefs, thoughts, feelings, motivational states, and ongoing patterns of thought, feeling, motivation, and reaction, including the relatively stable maintenance of an organization over time that includes an overall good grip on reality.

This, then, is one way things can look when a person has successfully integrated a false unconscious belief. It is not the only way, and it may be but a stage in a larger process. (p. 323) There is still the question of what motivations lie behind this belief. Moreover, though this resolution is well attuned to reality and stable so far as the particular conflict

is concerned, it may be threatened or undercut by other psychodynamic processes or co-opted for other purposes. Moreover, certain defensive structures can look superficially similar to this resolution; here the difference might be marked, for instance, by the superficiality of the irony, the lack of warm self-acceptance, the casual flimsiness of the patient's claim to know that he is not a thief, or the panic-concealing giddiness of his statement that he is one.

My use of the phrase 'speaking with one voice' emphasizes a connection with an important theme in recent work by Jonathan Lear. Lear has articulated a broadly neo-Aristotelian concept of integration involving excellent communication between the rational and non-rational aspects of the psyche, communication that enables a person to move forward freely to satisfy both the person's conscious aims and unconscious yearnings (Lear 2017). As he puts it:

> If the voices of the non-rational soul are an occasion for a creative, in-tune and thoughtful response from reason; and if, in turn, reason is able to enliven and free up the voices of the non-rational soul, as it channels them into a life worth living, we can give content to the thought that this is a rich form of speaking with the same voice.

> (Lear 2017: 49)

Here Lear calls attention to the way in which two systems or 'parts' of the psyche can work in harmony despite their different modes of functioning: they are, at bottom, in a certain kind of accord. This is an important form of integration. However, it differs from the mode of integration which I have been highlighting: the integration of conflicting mental states into a whole that neither evades nor elides the conflict but rather incorporates it into a broader, well-functioning, self-conscious standpoint.

Lear's own clinical examples do not thematize the relation between these two modes of integration. One of his clinical examples—the case of Mr B—is described in such a way that it appears that excellent communication is achieved without the patient's consciously integrating whatever conflict is holding him back (2017: 42). In the example of Ms A, by contrast, the patient self-consciously experiences the conflict for what it is, and this appears to play a role in enabling her to make a request that had been difficult for her (2017: 22). This is a moment that is comparable to what I have described in the integration of unconscious belief.

It is an open question for me to what extent and in what ways excellent communication between the rational and non-rational aspects of the psyche involves consciously experiencing and integrating conflicting mental states, just as it is an open question in what way(s) the latter form of integration might depend upon the form of good communication that Lear describes. The point that I want to mark for now is just this: because these phenomena are conceptually distinguishable, these questions need to be

considered. As Lear has helped us see, we shouldn't assume in advance that we know what integration is or should be.

(p. 324) Philosophical Implications: Psychic Unity, the First-Person Perspective, Transparency, and Rational Self-Constitution

I have argued that if we are to make sense of the possibility that bringing unconscious beliefs to consciousness could be a path to psychic health, we must make sense of the idea of self-consciously occupying the standpoint of a belief that one knows to be false. I have tried to show how to do so. In closing, I will briefly sketch some important implications for several large issues in philosophy.

a. There is a strong tendency amongst philosophers to view both the unity of the subject and psychic integration as matters of *rational* unity. For instance, in his defence of the broadly Freudian idea that the mind can contain partitions between distinct psychic systems, Donald Davidson comments, 'I postulate such a boundary somewhere between any (obviously) conflicting beliefs.... Such boundaries … are conceptual aids to the coherent description of genuine irrationalities' (Davidson 1986: 211; see also 1982). Such aids are needed, Davidson tells us, because 'two obviously opposed beliefs could coexist only if they were somehow kept separate, not allowed to be contemplated in a single glance': these are 'beliefs which, allowed into consciousness together, would destroy at least one' (Davidson 1997: 220). But the result of the partition, he says, is 'a single mind not fully integrated' (1997: 221).

These formulations presuppose a picture of psychic integration as *rational* integration, resulting in the rational abandonment of one or the other of the conflicting beliefs. The integrated mind is thus free from obvious conflict, on this picture. However, we have seen that self-conscious belief isn't tied that closely to rationality, nor does psychic integration necessarily involve the removal of rational conflict.[18]

(p. 325) Sydney Shoemaker deploys framing assumptions similar to Davidson's in considering Moore's paradox. He writes, of a person who affirms 'I believe that p but p is false', that 'the only way to save the coherence of this case is to suppose that it involves there being a divided mind. One part of a person's mind believes something, and another does not, and the part that does not believe it ascribes the belief to the part that does' (Shoemaker 2012: 249). Here Shoemaker gets things exactly backwards. The analysand's ability to sincerely say or think such a thing is a manifestation precisely of the overcoming of psychic division: it's from one and the same perspective—that of the conscious subject of the belief in question—that the analysand both self-ascribes the belief and acknowledges its falsehood.

At bottom, the lesson is this. There is a widespread view that it cannot possibly be correct to ascribe an obviously irrational combination of contemporaneous conscious attitudes to a person. Consideration of the clinical task of integrating false unconscious beliefs shows that this view is simply incorrect. The unity of a single, self-conscious, subjective perspective need not be a rational unity.[19]

b. I turn now to an issue prominent in recent philosophical discussions of knowledge of one's own beliefs and other attitudes. It is a familiar point in these discussions that there is a crucial difference between ordinary first-personal awareness of our emotions, beliefs, desires, and other attitudes, on the one hand, and, on the other hand, knowledge we might gain about these matters in other ways. For instance, someone might learn of her unconscious belief by accepting an interpretation offered by her analyst or by drawing an inference from indications in her behaviour and conscious thought and feelings. She doesn't thereby gain the sort of knowledge that we ordinarily have of our own beliefs. The difference is that her self-attribution is analogous to a third-personal psychological attribution: she does not self-ascribe the belief from the standpoint of the conscious subject of the very belief being self-ascribed. As it is put in the literature, she does not stand in a genuinely *first-personal* relation to her belief (Moran 2001).

According to a long and prominent philosophical tradition, the genuinely first-personal relation to one's attitudes is inextricably linked to considerations of rationality and a conception of the self as rational agent. Perhaps the most prominent proponent of such a view is Kant, who writes that 'it is only as intelligence that [a person] is his proper self' and that if certain desires cannot be viewed by the individual as the products of reason, 'He does not even hold himself responsible for these inclinations or impulses or attribute them to his proper self' (Kant *Grundlegung* 457–8; Beck's translation p. 7). Themes along these lines have shaped the thinking of a large number of contemporary philosophers, not all of whom identify as specifically Kantian (Bilgrami 2006; Boyle 2009; Moran 2001; Rödl 2007). For instance, some hold that self-attribution of beliefs from the first-person stance necessarily involves a stance of deliberative agency and rational endorsement. Others hold that in first-personal attributions of belief, one sees the belief as the product of one's rational faculties.

(p. 326) The points I have been urging show such views to be mistaken. Self-attribution from a first-person stance requires consciously occupying the attitude's subjective perspective on the world. In the cases I have been discussing this does not require or allow for rational endorsement of the belief or even seeing it as the product of the proper functioning of one's rational faculties. Nor in these cases does the first-person relation involve taking a stance of deliberative agency towards the belief, since one might very well recognize that one can't get rid of the belief through rational deliberation. Rather, in these cases a genuinely first-personal stance involves seeing the world from the subjective standpoint of the belief even while regarding it as false and as the product of non-rational belief-forming processes involving wish, phantasy, and defence. What is required here is acknowledgement and acceptance of the reality of one's psychic life, building on a certain kind of emotional openness to that which one finds repugnant or

unreasonable.[20] This is not reducible to any sort of intellectual exercise, nor is it compatible with identifying oneself (as subject) with one's rational capacities in particular.

Still, something is right in thinking that our reasoning capacities, and our understanding of ourselves as possessing and aiming to exercise such capacities, are relevant to our ability to take up a first-personal relation to our beliefs. Our ability to acknowledge and accept our false beliefs depends fundamentally on our capacities for rational reflection and response to reality. Without those capacities, we could not contain the conscious false belief within a broader conscious perspective that puts it in its place, but would rather be the helpless victims of unresolved conflict.

c. Problems also arise for the related accounts of first-personal self-knowledge which take 'transparency' to be central. Gareth Evans claimed that in self-ascribing a belief about whether there will be a third world war 'I must attend … to precisely the same outward phenomena as I would attend to if I were answering the question' whether there will be a third world war (1982: 225): that is, to determine whether I believe that p, I ask whether p and then if I find that p is true (or false), I append 'I believe that . . .' to that result. As Moran puts it, '[T]he claim [of the Transparency Condition] … is that a 1st-person present tense question about one's belief is [to be] answered by reference to (or consideration of) the same reasons that would justify an answer to the corresponding question about the world' (Moran 2001: 62). On this conception, present-tense self-attribution of belief from the first-person standpoint is dictated by rational consideration of the relevant evidence and facts concerning the matters the belief is about.

This view cannot make sense of the clinical phenomena. When an analysand integrates a formerly unconscious belief that is plainly false, she cannot self-attribute the belief by following the transparency procedure; after all, she recognizes that the belief is false, so when she asks whether p her answer will point away from self-attribution of belief that p —at least if she follows the transparency procedure. Still, she stands in a properly first-personal relation to her belief that p nonetheless. Her relation to that belief is not merely 'attributional' or 'theoretical'; she occupies the subjective perspective of the belief and (p. 327) self-attributes the belief from that position. Properly first-personal self-attribution thus need not conform to the transparency condition.

This point is of clinical importance. Precisely because the transparency procedure ties self-attribution to one's judgement about what is the case or what the relevant reasons support, to take up the transparency approach in the clinical situation would be to refuse to admit into one's self-conscious view any belief that one can see fails to fit with the relevant evidence or facts. The stance of Transparency would thus be a manifestation of resistance to the analytic process or a defence against conscious psychic conflict in these cases. In the clinical setting, the key question is, 'How am I oriented towards myself and my world?' If one is tempted to answer that question by focusing on the question, 'What is there reason for here?', one might well be led into a defensive, false self-conception (Bell and Leite 2016; Lear 2011: 56).

Some accounts of self-knowledge emphasize transparency without framing it in terms of the Kantian themes highlighted earlier. Clinical cases raise a challenge for these views as well.

For instance, according to Alex Byrne's (2005, 2011) 'inferentialist' account, first-personal awareness of our own beliefs is secured through an inference in accordance with the 'Doxastic Principle', 'If p, then conclude that you believe that p'. Byrne emphasizes that this would be a distinctively first-personal way of acquiring such knowledge; the inference's reliability depends precisely on the fact that it concerns *one's own* beliefs. However, the patient who comes to see 'from the inside' that she believes—quite incorrectly—that she has stolen everything good in life, has genuinely first-personal knowledge of this belief, and yet we cannot reconstruct this knowledge in terms of the reasoning Byrne proposes. First, the person would not reason in the way he suggests ('I have stolen everything good in my life, so I believe that I have stolen everything good in my life'), because she regards the premise as false. Second, it is hard to see how an inference from a premise a person correctly regards as false could give her knowledge of the conclusion. Her first-personal awareness of this belief consequently cannot be grounded in any such reasoning.

d. I turn to one last issue: the constitution of the self. A prominent strand of thought in contemporary philosophy urges that *my endorsement* is what makes the difference between that which belongs to 'me' as subject and that in my psychology to which I relate only as an object of description, report, and possibly management (Frankfurt 1988; Korsgaard 2009).[21] On this view my unity as a subject is threatened to the extent that I have not yet sorted conflicting elements into 'me' and 'not-me' through deliberation and endorsement or rejection. Korsgaard gives characteristically sharp expression to the point.

> Because human beings are self-conscious, we are conscious of threats to our psychic unity or integrity. Sometimes these threats spring from our own desires and impulses … we deliberate in the face of threats to our integrity, and as against them … we must repress them in order to be one, to be unified, to be whole (2009: 26).

(p. 328) Korsgaard does not merely mean that we give order to our psychologies by choosing to act on some motives and beliefs but not others. Rather, she thinks of deliberation and choice as forging a distinction between what is genuinely 'us' and what is not. 'That's what deliberation is: an attempt to reunite yourself behind some set of movements that will count as your own' (Korsgaard 2009: 213); through deliberation and choice we engage in 'the endorsement of our identities, our self-constitution' (2009: 43). So, for Korsgaard self-unity is achieved through extrusion: I cannot recognize from the first-person point of view that something belongs to me *as subject* which I do not and cannot endorse, but rather must regard it merely as 'a product of some force that is at work *on* me or *in* me' (2009: 18–19).

This gets things backwards from the clinical point of view. What is true, given a psychodynamic framework, is rather that psychic health requires us to *integrate* these threats 'in order to be one, to be unified, to be whole.' That doesn't mean that we must endorse them or choose to act upon them. It means that we must bring them within our purview as self-conscious subjects, understanding them as *us*, as part of our identity, as part of who we are.[22]

References

Arlow, J. (2008/1969). 'Unconscious fantasy and disturbances of conscious experience'. *Psychoanalytic Quarterly* 77(1): 21–46.

Auchincloss, E. L. and Samberg, E. (eds) (2012a). 'Fantasy'. In *Psychoanalytic Terms and Concepts*. New Haven, CT: Yale University Press.

Auchincloss, E. L. and Samberg, E. (eds) (2012b). 'Splitting'. In *Psychoanalytic Terms and Concepts*. New Haven, CT: Yale University Press.

Beebe, B. and Lachmann, F. (1994). 'Representation and internalization in infancy: Three principles of salience'. *Psychoanalytic Psychology 11*: 127–165.

Bell, D. and Leite, A. (2016). 'Experiential self-understanding'. *The International Journal of Psychoanalysis 97*(2): 305–332.

Bilgrami, A. (2006). *Self-Knowledge and Resentment*. Cambridge, MA: Harvard University Press.

Bohleber, W., Jiménez, J. P., Scarfone, D., Varvin, S., and Zysman, S. (2015). 'Unconscious phantasy and its conceptualizations: an attempt at conceptual integration'. *International Journal of Psychoanalysis 96*(3): 705–730.

Boyle, M. (2009). 'Two kinds of self-knowledge'. *Philosophy and Phenomenological Research 78*(1): 133–164.

(p. 329) Britton, R. (1995). 'Psychic reality and unconscious belief'. *The International Journal of Psychoanalysis 76*(1): 19–23.

Britton, R. (1998). *Belief and Imagination*. New York, NY: Routledge.

Byrne, A. (2005). 'Introspection'. *Philosophical Topics 33*: 79–104.

Byrne, A. (2011). 'Transparency, belief, intention'. *Proceedings of the Aristotelian Society* Supplementary Volume *85*: 201–221.

Child, W. (1996). *Causality, Interpretation, and the Mind*. Oxford: Oxford University Press.

Coliva, A. (2015). 'How to commit Moore's Paradox'. *Journal of Philosophy 112*(4): 169–192.

Davidson, D. (1970/1980). 'Mental acts'. In *Essays on Actions and Events* (pp. 207–225). Oxford: Oxford University Press.

Davidson, D. (1982/2004). 'Paradoxes of irrationality'. In *Problems of Rationality* (pp. 169–187). Oxford: Oxford University Press.

Davidson, D. (1983/2001). 'A coherence theory of truth and knowledge'. In *Subjective, Intersubjective, Objective* (pp. 137–153). Oxford: Oxford University Press.

Davidson, D. (1984). *Essays on Truth and Interpretation*. Oxford: Oxford University Press.

Davidson, D. (1986/2004). 'Deception and division'. In *Problems of Rationality* (pp. 199–212). Oxford: Oxford University Press.

Davidson, D. (1997/2004). 'Who is fooled?' In *Problems of Rationality* (pp. 213–230). Oxford: Oxford University Press.

Eagle, M. (2013). 'The implications of conceptual critiques and empirical research on unconscious processes for psychoanalytic theory'. *Psychoanalytic Review 100*(6): 881–917.

Erreich, A. (2003). 'A modest proposal: (re)defining unconscious fantasy'. *Psychoanalytic Quarterly 72*(3): 541–574.

Evans, G. (1982). *Varieties of Reference*. Oxford: Oxford University Press.

Feldman, M. (2007). 'Addressing parts of the self'. *International Journal of Psychoanalysis 88*(2): 371–386.

Fonagy, P. (1999). 'Memory and therapeutic action'. *International Journal of Psychoanalysis 80*(2): 215–223.

Frankfurt, H. (1988). 'Freedom of the will and the concept of a person'. In *The Importance of What We Care About* (pp. 11–25). Cambridge: Cambridge University Press.

Freud, S. (1915). 'The unconscious'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud,* Vol. 14 (pp. 159–215). London: Hogarth Press.

Freud, S. (1917). *Introductory Lectures on Psychoanalysis*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud,* Vol. 15. London: Hogarth Press.

Freud, S. (1927). 'Fetishism'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 21 (pp. 152–157). London: Hogarth Press.

Freud, S. (1940a). 'An outline of psychoanalysis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 23 (pp. 144–208). London: Hogarth Press.

Freud, S. (1940b). 'Splitting of the ego in the process of defence'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud,* Vol. 23 (pp. 275–278). London: Hogarth Press.

Gardner, S. (1993). *Irrationality and the Philosophy of Psychoanalysis.* Cambridge: Cambridge University Press.

Gendler, T. S. (2008). 'Alief and belief'. *Journal of Philosophy 105*(10): 634–663.

(p. 330) Gertler, B. (2011). 'Self-knowledge and the transparency of belief'. In Anthony Hatzimoysis (ed.), *Self-Knowledge* (pp. 125–145). Oxford: Oxford University Press.

Heal, J. (1994). 'Moore's paradox: a Wittgensteinian approach'. *Mind 103*(409): 5–24.

Hursthouse, R. (1991). 'Arational actions'. *Journal of Philosophy 88*(2): 57–68.

Isaacs, S. (1948). 'The nature and function of phantasy'. *International Journal of Psychoanalysis 29*: 73–97.

Kant, I. (1990). *Foundations of the Metaphysics of Morals*, trans. L. W. Beck. New York: Macmillan.

Kernberg, O. (1966). 'Structural derivatives of object relationships'. *International Journal of Psychoanalysis 47*: 236–252.

Klein, M. (1946). 'Some notes on some schizoid mechanisms'. In *Collected Works of Melanie Klein*, Vol. 3: *Envy and Gratitude and Other Works* (pp. 1–24). London: Hogarth Press.

Korsgaard, C. (2009). *Self-Constitution: Agency, Identity, and Integrity*. Oxford: Oxford University Press.

Lear, J. (2003). *Therapeutic Action*. New York: Other Press.

Lear, J. (2011). *A Case for Irony*. Cambridge, MA: Harvard University Press.

Lear, J. (2017). *Wisdom Won From Illness*. Cambridge, MA: Harvard University Press.

Leite, A. (2016). 'Changing one's mind'. *Philosophy and Phenomenological Research*, version of record online 7 Nov. 2016. doi:10.1111/phpr.12332

Leite, A. (2018). 'Expressivism, psychic health, and the first-person stance'. *Journal of Personality Disorder* (special issue on self-knowledge and personality disorder) *32*(3): 351–373.

Lewis, D. (1982). 'Logic for equivocators'. Noûs *16*(3): 431–441.

Lyons-Ruth, K. (1999). 'The two-person unconscious: Intersubjective dialogue, implicit relational knowing, and the articulation of meaning'. *Psychoanalytic Inquiry 19*: 567–617.

Mölder, B. (2010). *Mind Ascribed: An Elaboration and Defense of Interpretivism*. Philadelphia, PA: John Benjamins.

Moore, G. E. (1993). 'Moore's Paradox'. In T. Baldwin (ed.), *G. E. Moore: Selected Writings* (pp. 207–212). London: Routledge.

Moran, R. (2001). *Authority and Estrangement*. Princeton, NJ: Princeton University Press.

Parrott, M. (2015). 'Expressing first-person authority'. *Philosophical Studies 172*: 2215–2237.

Rödl, S. (2007). *Self-Consciousness*. Cambridge, MA: Harvard University Press.

Schwitzgebel, E. (2002). 'A phenomenal dispositional account of belief'. *Noûs 36*(2): 249–275.

Shoemaker, S. (2012). 'Self-intimation and second-order belief'. In D. Smithies and D. Stoljar (eds), *Introspection and Consciousness* (pp. 239–258). Oxford: Oxford University Press.

Steiner, J. (1993). *Psychic Retreats: Pathological Organizations in Psychotic, Neurotic and Borderline Patients*. London: Routledge.

Stern, D. N., Sander, L. W., Nahum, J. P., Harrison, A. M., Lyons-Ruth, K., Morgan, A. C., Bruschweilerstern, N., and Tronick, E. Z. (1998). 'Non-interpretive mechanisms in psychoanalytic therapy: the "something more" than interpretation'. *International Journal of Psychoanalysis 79*: 903–921.

Williams, P. (1996). 'Panel report—psychic reality: theoretical concepts'. *International Journal of Psychoanalysis 77*: 79–84.

Wittgenstein, L. (2001/1953). *Philosophical Investigations*. Oxford: Blackwell Publishing.

Wollheim, R. (2003). 'On the Freudian unconscious'. *Proceedings and Addresses of the American Philosophical Association 77*(2): 23–35.

Notes:

([1]) Schwitzgebel (2002) offers a broadly dispositional approach to belief that fits well with the perspective sketched here.

([2]) There are two options here because someone could have patterns and dispositions of conscious response—and be aware of the world in ways—that constitute occupying the subjective perspective of the attitude, and yet be unable to self-ascribe it. For instance, she might lack facility with relevant mental state concepts. More subtly, to consider just one sort of alternative in which she possesses facility with the relevant concepts, she

might consciously experience reactions of anger or envy but for defensive reasons be unable to experience them *as* anger or as envy. Finkelstein (this volume) offers an expressivist account of the conscious/unconscious distinction which attempts to cover all of the relevant cases. However, it is doubtful that his account adequately handles clinical examples in which a person directly expresses her attitude through self-attribution and yet fails to understand in the requisite way what she is doing (Bell and Leite 2016; Leite 2018). More complicated versions of such examples cause trouble for Finkelstein's account even when it is supplemented with the requirement (Finkelstein, this volume) that the person be able to 'co-expressively gloss' the expression of the attitude.

(^{3}) Only a great deal of experience with the person will enable a belief attribution of this specificity. For instance, some of the data I have described is compatible with an attribution of the belief 'I am an imposter', but someone who thinks himself an imposter rather than a thief will exhibit a different profile of thoughts and emotional responses. However, some measure of indeterminacy may be an inevitable part of mental state attributions as well. (Nothing I have said here addresses questions of clinical technique such as how the analyst should regard her own conjectures about the patient or how she should bring them into the clinical exchange.)

(^{4}) This is a schematic characterization of the mechanism of 'disavowal' proposed by Freud in his attempt to account for certain cases of fetishism (1927, 1940a: 201–4, 1940b).

(^{5}) The Kleinian tradition uses the spelling 'phantasy' to mark off the specifically psychoanalytic concept, as do many other analysts writing in Britain. However, many non-Kleinian psychoanalysts, especially those in America, prefer the spelling 'fantasy'. For convenience, I will follow the Kleinian spelling. No tendentious theoretical assumptions are thereby intended.

(^{6}) This difficult task often isn't helped by saying that a part of the patient hates the object while another part loves the object. As Michael Feldman puts it, 'while the patient might be able to understand such interpretations intellectually, [she] may not be able to use it in an insightful way, to promote psychic integration' (2007: 383).

(^{7}) Not all cases are of this sort; for instance, many repressed beliefs in relation to trauma are not false.

(^{8}) I don't deny that sometimes we simply don't know what to say about what the person believes or knows. However, the case relevant here isn't like that. It is one in which the person has the dispositions (described earlier) involved in unconsciously believing that he is a thief and *also* has the dispositions involved in believing—indeed, knowing—that he isn't one. The complex interplay between these dispositions explains the variegated pattern of actual behaviour and reactions.

(^{9}) The integration of unconscious belief thus poses a particularly sharp challenge for this philosophical tradition. It has been noted in recent work on Moore's paradox that in

situations of irrational or superstitious belief, a person might recognize that she has a belief even while acknowledging its falsity (Gertler 2011; discussion in Coliva 2015). However, it hasn't been clear in these examples whether the person occupies the subjective position of the belief at the time of the self-attribution. If not, then such examples are beside the point. Just as it is no surprise that a person could recognize on the basis of her own behaviour that she has a false unconscious belief, it is no surprise that self-attribution is straightforward when one is alienated from one's belief and does not at that moment occupy its subjective standpoint on the world. This is why the integration of unconscious false belief is a key test case.

(^{10}) This phenomenology is distinct from *what* is presented as being the case, but rather concerns the manner in which it is so presented.

(^{11}) It is sometimes asserted that understanding belief as involving 'commitment'—say, a 'commitment' to the truth of the believed proposition—precludes intelligible self-ascription of false belief (Coliva 2015). However, on any ordinary conception of commitment you can perfectly well recognize that you have and are expressing a commitment that you shouldn't have: a commitment that is erroneous, mistaken, or misguided. Admittedly, 'commitment' could be stipulatively defined in such a way that 'commitment' to the truth of p is incompatible with the self-attribution of false belief. However, one had better not understand conscious belief in such terms. Such an account has the consequence that self-conscious, deliberative change of mind is impossible—which it surely isn't (Leite 2016).

(^{12}) Thanks here to Sarah Majid.

(^{13}) Another sort of short circuit would arise if the analysand defensively rolled up his sleeves and got to work on trying to change the belief before truly coming to occupy it (Wollheim 2003).

(^{14}) Fonagy, for instance, writes: 'Therapeutic action lies in the conscious elaboration of preconscious relationship representations' (1999: 218). Stern et al. comment: 'In the course of an analysis some of the implicit relational knowledge will get slowly and painstakingly transcribed into conscious explicit knowledge' (1998: 918). Lyons-Ruth writes: 'If relational knowing is as much implicit and procedural as symbolic, the work of elaborating new implicit procedures for being with others must occur at enactive *as well as* symbolic levels … For an adult patient, more collaborative and inclusive dialogue may involve partially translating previously implicit procedural knowing into words' (1999: 608, 611, italics added).

(^{15}) This example is suggested by a discussion in Lyons-Ruth (1999: 596).

(^{16}) Even Stern et al. recognize the clinical importance—and difficulty—of bringing dynamically unconscious beliefs to consciousness. They write: 'The process of rendering repressed knowledge conscious is quite different from that of rendering implicit knowing

conscious. They require different conceptualisations. They may also require different clinical procedures, which has important technical implications' (1998: 918).

([17]) There is a link here—which only became clear to me after writing this chapter—to Lear's important work on irony and therapeutic action (2003: esp. 118–33; see also Lear 2011).

([18]) Psychoanalysts, too, are sometimes tempted by the conception of integration as an absence of conflict. For instance, Morris Eagle writes, 'In his later writings, [Freud] writes that in health there is a "seamlessness" between the id and the ego, and also that the overarching goal of psychoanalysis is unity of the personality. In taking this general perspective, Freud aligns himself with those whose philosophical and spiritual ideal is being at one with oneself, a man or woman of one piece. These figures include Confucius … who reported that at the age of seventy … the dictates of his heart and his sense of right and wrong were now one and the same, and Kierkegaard … who wrote that purity of heart is to will one thing. The primary value underlying this idea is … the experience of being at peace with oneself.… One way of putting it is to say that psychoanalysis moves a bit out of what appeared to be the near-exclusive orbit of the Enlightenment vision and closer to philosophical traditions characterized by a central concern with the achievement of inner harmony' (Eagle 2013: 908). Here we see a slide from 'seamlessness' (which is a functional/structural concept) to an ideal of integration as the removal of conflict. However, seamlessness and the unity of the personality (understood in functional/structural terms) do not in fact require the removal of conflict, but rather acceptance of the conflict—of both sides of the conflict as 'me'—and the discovery or creation of a way to move forward given the fact of conflict. This alternative ideal—of living with conflict gracefully and in full consciousness—gives us a different conception of what 'being at peace with oneself' and 'inner harmony' (Eagle 2013: 908) might come to.

([19]) I argue for this point on other grounds as well in Leite 2016.

([20]) There is a parallel issue that can arise regarding psychoanalytic empathy—the analyst's capacity to empathically share in the patient's perspective without endorsing it and while maintaining her own.

([21]) Of course, this characterization overlooks the vast differences between Frankfurt's and Korsgaard's views in other respects.

([22]) I am grateful to many people over many, many years for assistance related to this chapter. I would especially like to thank Kate Abramson, David Bell, David Bleecker, Matt Boyle, Louise Braddock, Jason Bridges, Ronald Britton, Fred Busch, Morris Eagle, Gary Ebbs, Peter Fonagy, Richard Gipps, Louise Gyler, Edward Harcourt, Jim Hopkins, Leon Kleimberg, Michael Lacewing, Jonathan Lear, Richard Moran, Caroline Polmear, and Mary Target. I apologize to the many additional people who I am surely forgetting. I would also like to thank the audiences at the 2013 Auburn University Conference on Theoretical Rationality, the 2014 University of London/London Institute of Psychoanalysis Joint Conference on Philosophy and Psychoanalysis, the 2015 American Philosophical

Association Pacific Division Meeting, and the members of the Philosophy and Psychoanalysis Study Group of the London Institute of Psychoanalysis.

Adam Leite

Department of Philosophy, Indiana University, Bloomington, USA

Oxford Handbooks Online

Abstract and Keywords

How should we understand the distinction between conscious states of mind and unconscious ones? This chapter briefly reviews an answer to this question that the author has set out and defended in earlier work; it then suggests a new answer—one that supplements, rather than replaces, the old answer. In spelling out this new answer, the chapter offers an account of a distinction that is related to, but not identical with, that between conscious and unconscious states of mind, viz. the distinction between conscious and unconscious expressions.

Keywords: expression, expressivism, conscious, consciousness, first-person authority, self-knowledge, unconscious

1. A Question

How should we understand the distinction between conscious states of mind and unconscious ones? This question, which I'll refer to as 'Q' in what follows, could also be put this way: by virtue of what is someone's anger, fear, anxiety, or desire, for example, rightly characterized as either conscious or unconscious? In what follows, I'll introduce some background to, and then briefly review, an answer to Q that I've set out and defended in earlier work.[1] I'll raise a worry about how this answer could be correct, and I'll discuss an exchange between Jonathan Lear and Richard Moran that bears on how the answer I've given to Q might help to elucidate what Lear calls 'the psychoanalytic meaning of making the unconscious conscious' (Lear 2011: 52).[2] Both the worry and the exchange between Lear and Moran will lead me in the direction of a new answer to Q— i.e. toward another way of understanding the difference between conscious and unconscious states of mind—one that supplements, rather than replaces or corrects, the old answer. Providing it will require that I present an account of a distinction that is related to, but not identical with, that between conscious and unconscious states of mind, viz. the distinction between (what I'll call) conscious and unconscious expressions.

Making the Unconscious Conscious

A good point of departure for theorizing about what it means to characterize someone's state of mind as either 'conscious' or 'unconscious' is a simple, tempting, and ultimately unsatisfactory answer to Q that can be put as follows: 'A person's mental state is conscious if she is aware of it, i.e. if she knows that she is in it. It is unconscious if she does not know that she is in it'. To see why this is not a satisfactory answer to Q, consider the following story:

> Max is a philosophy professor who lives and works in Chicago. Every few months, his mother flies to Chicago from her home in New York to meet with business associates and to visit her only child. Almost every time she does this, Max forgets to pick her up at the airport. Now, as a rule, Max doesn't tend toward this sort of forgetfulness. He never forgets to retrieve his wife or his friends from the airport. Moreover, on those occasions when he forgets to retrieve his mother, he also manages to be away from his mobile phone or to be carrying a phone whose battery has died—so, she is unable to reach him. One day, while talking with a colleague about his mother, it occurs to Max that this behaviour of his might constitute evidence that he is unconsciously angry at her. Indeed, on thinking through some of his other recent behaviour involving his mother, Max becomes convinced by this hypothesis—convinced on the basis of behavioural evidence that he harbours unconscious anger toward her. He says to his colleague, 'I must be unconsciously angry. It's the only way to explain the way I've been acting; I'm angry at my mother'.

Imagine that Max is justified in drawing the conclusion that he does. In such a case, we might describe him as 'knowing that he is unconsciously angry at his mother'. I take it that there is nothing incoherent in such a description. But if this is so, it means that being consciously angry is not the same as knowing that one is angry. At the end of the story, Max knows that he is angry at his mother; he is aware of his anger, but this anger is still unconscious.

Why is it tempting to think that a conscious mental state is just one that its subject knows about or is aware of? I suspect that part of the answer lies in the fact that we use the expression 'conscious of' (nearly) interchangeably with 'aware of'. Right now, as I write these words, I'm aware of—i.e. conscious of—my dog's left front paw, which is resting on my right foot. Of course, no one imagines that it somehow follows from the fact that I'm conscious *of* Kopi's paw that her paw can itself be described as 'conscious'. But the difference between 'conscious *of* x' and 'x's *being* conscious' is easier to get confused about when x is one's own state of mind. In order to get ourselves into a position from which we can think clearly about question Q, we need to distinguish between two ways in which the word 'conscious' is used. At the conclusion of Max's story, he has become aware of—we could say 'conscious of'—his anger toward his mother. But his anger is not conscious. He is conscious *of* his anger, but he is not conscious*ly* angry. (It is only when someone is conscious*ly* angry—and not merely conscious of his own anger—that his anger is said to *be* conscious). Q asks how we should understand the difference between conscious states of mind and unconscious ones (i.e. the difference between someone's

being consciously, e.g. angry and his being unconsciously angry). We'll fail to get this difference into focus if we conflate it with a distinction between things that one is aware of and things that one is unaware of.[3]

2. Wittgenstein's Suggestion

The answer to Q that I've defended in the past may be understood as an attempt to develop a suggestion that Ludwig Wittgenstein makes concerning how to think about psychological self-ascriptions. In his late writings, Wittgenstein often suggests that we should understand such ascriptions as, or as akin to, *expressions*. Thus he writes:

> The statement "I am expecting a bang at any moment" is an *expression* of expectation.
>
> (Wittgenstein 1967: §53)
>
> When someone says "I hope he'll come"—is this a *report* about his state of mind, or an *expression*[4] of his hope?—I can, for example, say it to myself. And surely I am not giving myself a report.
>
> (Wittgenstein 1953: §585)
>
> For even when I myself say "I was a little irritated about him"—how do I know how to apply these words so precisely? Is it really so clear? Well, they are simply an expression.
>
> (Wittgenstein 1992: 70)

Wittgenstein thinks that, at least when we're doing philosophy, we are inclined to imagine that saying, 'I expect an explosion', is more like saying, 'He expects an explosion', than in fact it is. Thus, we assume that just as my ascribing an expectation of an explosion to another person requires that I have some epistemic ground or justification for believing that he expects an explosion, so, too, my ascribing an expectation of an explosion to *myself* requires that I have some epistemic ground or justification for believing that I expect an explosion. The likely result will be a view according to which, even under the best of circumstances (even when there is no question of my being opaque to myself), psychological self-ascriptions should be understood as reports of observations—a view that Wittgenstein takes to be misguided.[5] His suggestion—that we think of such ascriptions as expressions—is meant to provide an alternative way of understanding their grammar. When Wittgenstein calls, e.g. a self-ascription of expectation an 'expression', part of what he means is that it is not the report of an observation. Thus, there is a fundamental grammatical asymmetry between psychological self-ascriptions and ascriptions of psychological states and goings-on to others.

In Wittgenstein's *Philosophical Investigations*, this suggestion first appears in §244, in the context of an exchange between Wittgenstein and his interlocutor concerning how it is

that children manage to learn the names of sensations. During this exchange, we find Wittgenstein saying:

> Here is one possibility: words are connected with the primitive, the natural, expressions of the sensation and used in their place. A child has hurt himself and he cries; and then adults talk to him and teach him exclamations and, later, sentences. They teach the child new pain-behaviour.

> (Wittgenstein 1953: §244)

It's no accident that it is in the context of a discussion of language learning that Wittgenstein introduces the idea that some self-ascriptions might be understood in terms of expression. Part of his point could be put as follows (and here I'm reading §244 through the lens of many other passages in which Wittgenstein likens self-ascriptions to expressions). No one imagines that when an infant expresses a pain by crying, wincing, or moaning, he is issuing a report concerning an item that he has discovered via some sort of inwardly directed observation. We can think of a child's gaining the ability to say, 'Ouch!' or, 'That hurts!' or, 'I feel a pain in my foot', as a matter of his learning new ways to do something that, in some sense, he could already do before he could talk, viz. express his pains. And while an expression that takes the form of a prelinguistic cry or wince *is* different in significant ways from an expression that takes the form of a self-ascription[6]—even so, Wittgenstein thinks, there is no good reason to hold that when a child becomes able to produce the latter sort of expression, he suddenly needs to rely on inner observation in order to manage it.

Immediately following the bit of *Investigations* §244 that I quoted in the preceding paragraph, the text continues as follows:

> "So you are saying that the word 'pain' really means crying?"—On the contrary: the verbal expression [*der Wortausdruck*] of pain replaces crying and does not describe it.

The question that Wittgenstein's interlocutor asks in the first of these two sentences is liable to seem oddly confused or unmotivated. After all, Wittgenstein has not suggested (in §244 or elsewhere) that the word 'pain' *means* something different from what we always took it to mean; he's not suggested that it really refers to behaviour. What is moving the interlocutor to ask what he asks here?

Wittgenstein's interlocutor is in the grip of a false dilemma concerning sensations and other mental items. He (the interlocutor) consistently thinks—not only in the *Investigations*, but in other late writings as well—that one must choose between (1) thinking of pains, etc. as things that can be known about only via a kind of inner observation directed at a realm of objects that just one subject can, in principle, have direct access to, and (2) embracing a kind of behaviourism that, in effect, *eliminates* sensations and replaces them with behaviour. In the quoted passage from §244, the interlocutor gathers that Wittgenstein means to be rejecting the idea that we ordinarily

self-ascribe pains on the basis of inner observation, so he, in effect, accuses him of embracing behaviourism. In this instance, the interlocutor's *way* of accusing Wittgenstein of embracing behaviourism is to suggest that Wittgenstein thinks the word 'pain' doesn't after all pick out anything genuinely inner, but instead just means crying.

This isn't the only moment in the *Investigations* when Wittgenstein's interlocutor accuses him of behaviourism. But in §244, Wittgenstein replies to the accusation in a way that is liable to sound off-key. He doesn't say merely that an utterance such as 'I'm in pain' should be understood as a verbal expression of pain and not as a description of crying (or of any other behaviour). He says that the verbal expression of pain *replaces* ('ersetzt') crying. And perhaps this doesn't seem quite right. After all, it's not as if children *stop* expressing pains by crying when they learn how to express them by talking about them.

Maybe Wittgenstein ought to have used a different word at the end of §244. Still, I think it's important that there is, after all, a *kind* of replacement that occurs when children learn to express their pains by self-ascribing them. And I think Wittgenstein *might* have had this in mind when he imagined the exchange in the way that he did. In any case, *I* want to say that there is an important sense in which the *kind* of crying a child did as a prelinguistic infant really is no longer a part of her life after she learns to talk about her own sensations, emotions, and attitudes. The kind of crying she did as an infant is, at that point, replaced by another kind of crying. After she learns to talk, even though her expressions-of-pain-via-crying sometimes look and sound the same as they did before, they now could be said to have a different *form*. I'm going to have to leave this suggestion dark for a little while—until after I've discussed the possibility of an expression's being conscious. I'll come back to it in §7.

3. An Answer to Q

In my own writing, I've tried to take what Wittgenstein says concerning the grammar of our talk about our own mental states and use it to make sense of the distinction between two kinds of mental state. The guiding idea could be put this way: once a human being has acquired language, her mental states are such that she is, ordinarily, able to express them via the sort of speech act that Wittgenstein talks about; she can express her sadness, anger, or desire by ascribing it to herself. Insofar as she is able to do this, her sadness, anger, or desire is said to be *conscious*. But sometimes, even after a human being has acquired language (and with it, the capacity to say what she is thinking or feeling), she is unable to express, e.g. some particular fear of hers by self-ascribing it. A kind of expressive ability that is characteristic of normal psychological states and events as they figure in the life of a linguistic animal is, in such a case, absent or blocked. The word that we use to characterize this sort of absence is 'unconscious' or 'unconsciously'; we say that someone's fear is unconscious or that she is unconsciously afraid.

What would it take for Max's unconscious anger toward his mother to *become* conscious? It would not be sufficient for Max to become aware on the basis of behavioural evidence that he was angry. For Max's anger to become conscious, he would need to acquire, or to

regain, an ability to express it merely by self-ascribing it. This expressive ability is what he still lacks at the conclusion of the story I told about him in §1. At the end of that story, Max says to a colleague, 'It's the only way to explain the way I've been acting; I'm angry at my mother'. In so saying, Max expresses a belief (or perhaps knowledge) that he has about himself. But he does not express his anger at his mother. He may express that anger by stranding his mother at the airport, but he is still unable to express it by self-ascribing it. Hence, it is still unconscious.[7] This is to say, I've answered Q by claiming that what distinguishes conscious states of mind and unconscious ones is the presence or absence of a particular expressive ability.

In my *Expression and the Inner*, I put the point as follows:

> *Someone's mental state is conscious if he has an ability to express it merely by self-ascribing it. If he lacks such an ability with respect to one of his mental states, it is unconscious.*

> (Finkelstein 2003: 120)

What work is the word 'merely' doing in this encapsulation of an answer to Q? Why not leave out 'merely' and say, more simply, that a person's mental state is conscious if and only if he can express it in a self-ascription? The trouble with the simpler formulation is that it is open to counter-examples of the following sort. Imagine that Max occasionally expresses his anger by speaking in a peculiar, clipped tone of voice. *While speaking in this tone of voice*, he says: 'The only way to explain my odd behaviour is to posit that I am unconsciously angry at my mother. So, I'm angry at her'. Through his tone of voice, Max expresses his anger in a self-ascription of it. Even so, his anger is unconscious. In *Expression and the Inner*, I wrote:

> When I say that someone's state of mind is conscious if he has an ability to express it *merely* by ascribing it to himself, I mean this: the sort of ability at issue is one that enables a person to express his state of mind in a self-ascription of it, where what matters [to the self-ascription's expressing what it does] isn't his tone of voice (or whether he is tapping his foot, or what he is wearing, or to whom he happens to be speaking), but simply the fact that he is giving voice to his sincere judgment about his own state of mind.[8] That I might manage to express my unconscious anger in a self-ascription of it *via a clipped tone of voice* doesn't show that I have the relevant sort of expressive ability. When I am consciously angry, I can say in a neutral tone of voice, "I'm angry," and thereby express my anger.

> (Finkelstein 2003: 120)

Was that an adequate response to the tone-of-voice case? In one sense, I think it was. It showed that the case constitutes a merely apparent, not a genuine, counter-example to the view I was defending in the book (and that I still take to be correct). On the other hand, it seems to me now that what I wrote there (and elsewhere) is liable to leave a reader with a nagging question that might be put as follows: 'Why should it *matter* if a

person is able to express his anger merely by self-ascribing it—as opposed to being able to express it in a self-ascription that comes in a clipped tone of voice? How *could* the difference between conscious and unconscious anger come down to something so apparently unimportant as that?' I take the challenge posed by this question seriously. Indeed, part of what has moved me to pursue a second answer to Q is the aim of coming to better understand how, given this sort of challenge, my first answer could be correct. How is it that gaining, or regaining, an ability to express some emotion or attitude merely by self-ascribing it can make a significant difference in a person's life? I'm not yet in a position to address this question, but I'll return to it after I've given a second answer to Q, in the final section of this chapter.

4. Non-Linguistic Animals

Given the answer to Q that I've just been reviewing, we should not expect the distinction between conscious and unconscious psychological states to, as it were, *get a grip* when we are thinking about non-linguistic animals. After all, it is only when one is considering the psychology of a linguistic creature that it makes sense to distinguish between states that can be expressed in self-ascriptions and states that cannot. And, as a matter fact, we *don't* characterize the psychological states of non-linguistic creatures as either conscious or unconscious. Even those of us who are prone to attribute complex thoughts and feelings to our pets do not say things like: 'My dog gets upset when I pack my suitcase because she has an unconscious fear of being abandoned', or, 'Over time, my cat's distaste for the people who live next door has gone from being unconscious to being conscious'. Nor do we say such things about prelinguistic children. To borrow a phrase that Wittgenstein uses in a related context,[9] when a child learns to attribute mental states to herself, a 'new joint' is added to the language-game. We could say that when children learn to self-ascribe their psychological states and events, the *form* of those states and events changes. From then on, they are either conscious, unconscious, or somewhere in between conscious and unconscious.[10] To the extent that some desire of mine is such that I am unable to express it by self-ascribing it, it could be said to suffer from a formal defect or imperfection. But it is no imperfection in Kopi's desire, e.g. for water that she is unable to express it by self-ascribing it. Kopi's desires are, unlike mine, neither conscious nor unconscious.[11] This is to say: it makes no sense to describe a desire of hers as either 'conscious' or 'unconscious'. In the shift from brute to fully realized human desire, a new joint is added to the language-game; the psychological is transformed.

5. 'What is gained … by replacing some words for some tears?'

One of the goals that Jonathan Lear sets for himself in his fascinating *A Case for Irony* is to elucidate 'the psychoanalytic meaning of making the unconscious conscious' (Lear 2011: 52). In pursuing this end, he draws on the strand in my work that I've been reviewing and discussing in this chapter. Lear contrasts a case in which someone

concludes on the basis of self-observation that he must be angry, even though he isn't 'consciously feeling angry', with a very different case that he describes as follows:

> [I]n the midst of a boiling rage, I say "I'm *furious* with you!" In this case, the anger itself is present in the verbal self-ascription of anger that is directed at you. In this case, the self-ascription *of* anger is itself an *angry* expression. Unlike the former case, I do not have to observe myself to know that I am angry. I just *am angry*; and the form my anger at you takes on this occasion is the angry verbal self-ascription directed at you. In this case, the utterance "I am *furious* with you!" may replace other forms of angry expression. … This is a case, in Finkelstein's terms, in which I am not only conscious *of* my anger; I am *consciously angry*.
>
> (Lear 2011: 52–3)

Lear goes on to consider the question of how it might be therapeutic, rather than merely disruptive, to find some way of, as it were, taking unconscious anger—anger that has been 'held out of consciousness not simply because it is painful, but because it violates one's sense of who one is' (Lear 2011: 54)—and making it conscious. In this context, he says, 'The question thus becomes whether there could be a process—not *too* disruptive—by which the verbal expression of anger *replaces* (in the Wittgenstein-Finkelstein sense) its unconscious manifestations' (Lear 2011: 54).

A Case for Irony comprises a pair of lectures by Lear followed by responses by several other philosophers. In one of these responses, Richard Moran writes:

> The idea of "replacing" one mode of expression for another is not perfectly clear to me, particularly in the therapeutic context into which Lear is importing the idea. On a basic level, we might ask: If the two modes of expression are really doing the same work, then what is the *point* of "replacing" one with the other? If we take the notion of replacement literally, then what is gained, for instance, by replacing some words for some tears? Does that mean that the tears are now unnecessary, having been replaced by something else?
>
> (Moran 2011: 113)

It might be tempting to try to answer Moran's question by drawing a distinction between the kind of expressive 'work' that gets done via a self-ascription of sadness and the kind that gets done via crying. But I believe that this would be to draw a line in the wrong place. At least insofar as we're interested in psychoanalysis, there *is* an important line for us to draw between expressions that are, in a very particular way, defective or imperfect and ones that are not. But, as will become apparent, crying can lie on *either* side of this line. Moran's question is useful, I think, insofar as it helps us to see that we need to distinguish cases in which someone's (e.g.) crying constitutes (what can be called) a conscious expression of sadness (or joy or whatever) from ones in which someone unconsciously expresses her (e.g.) sadness via crying. Moran asks what the point is of replacing some tears with some words—that is, of replacing an expression that takes the

form of crying with one that takes the form of self-ascription. My own view is that there need not be *any* point in replacing crying with self-ascription. Tears are just fine as long as they constitute *conscious* expression. Psychoanalysis does not aim to replace tears with words. It aims to replace unconscious expressions with conscious ones.

6. The Distinction between Conscious and Unconscious Expressions

What does an unconscious expression look like? We already saw an example of one in §1, but I'd like to consider a somewhat different example now:

> One afternoon, Max and his wife, Sarah, have a quarrel while assembling hors d'oeuvres for a dinner party that they'll be hosting that evening. Sarah's new boss will be one of their guests, and partly for this reason, she is worried about the party's going well. Although the quarrel subsides after a few minutes, Max continues to dwell on it. Just before guests are scheduled to arrive, he reminds himself that Sarah has been anxious about the party's going smoothly, and he tells himself that it would be petty of him to express the anger that he's still feeling toward her in front of their guests. He resolves to be the perfect co-host, to be cheerful, charming, and useful. In spite of—and perhaps partly because of—this resolution, Max repeatedly expresses his anger during the party, but he remains unaware of doing so. He responds to Sarah's sharp, witty, and critical assessment of a recent movie by suggesting that she simply failed to appreciate what was innovative about it; he makes disparaging jokes about food that Sarah prepared; he tells a story about her brother that she finds embarrassing; and so on. It does not occur to Max that these remarks of his might be connected to that afternoon's quarrel, and if someone were to suggest that there *was* a connection, he would deny it. Nonetheless, it is apparent to all but the drunkest of their friends that Max is annoyed with Sarah. For all that, Max performs the bulk of the serving and most of the straightening up. Indeed, at one point during the evening, he privately congratulates himself on managing to 'put on a good face' for Sarah's sake, in spite of his angry, hurt feelings.

Here, Max unconsciously expresses his anger toward Sarah in passive-aggressive remarks. A noteworthy feature of this particular example of unconscious expression—a feature that is not present in every example—is that here, a person unconsciously expresses a state of mind that is, itself, conscious. (As I've asked you to imagine him, Max is consciously angry at Sarah). This example stands in contrast to one in which someone unconsciously expresses an unconscious state of mind. The story that I told in §1—in which Max expresses his unconscious anger by repeatedly stranding his mother at the airport—constitutes one such example.

Thus far, I've recounted two stories in which Max *unconsciously* expresses anger. Here's one in which he *consciously* expresses pleasure:

Early one evening, Max and Sarah are feeling both hungry and tired, so rather than cooking dinner, they decide to go to a nearby restaurant. Five minutes after they order, Sarah tells a joke, and after a brief pause, Max's face lights up in a smile. Sarah says, 'Are you smiling at my joke?' Max replies, 'No. I thought your joke *was* funny, but I'm smiling because I see the waiter with our food, and I'm really pleased that it's coming out so quickly'.

In this example, Max is able to speak to the question of what his smile expresses in, we could say, a first-personal way. His smile is a conscious expression of his pleasure at the imminent arrival of dinner.

How is conscious expression possible? How is it that Max is able to just say what his smile expresses? How does he know that he's *not* smiling at Sarah's joke? I want to approach these questions by first considering what seems to me an easier sort of case, one in which someone speaks with a kind of first-person authority—not about what psychological state a bit of his behaviour expresses, but rather—about what some of his own words mean:

One afternoon, Max and Sarah attend a philosophy department social event. A graduate student says to Max: 'I just met Sarah. Is she a philosopher too?' Max replies: 'She *is* a philosopher, by which I mean: she is moved by, and thinks hard about, philosophical questions. But she's not someone who makes her living by doing philosophy'.

Here, Max first says, 'Sarah *is* a philosopher', and then he provides a gloss on what he meant by these words. Imagine that someone were to ask: 'How did Max find out what he meant by the words, "Sarah *is* a philosopher"? Did he first hear himself utter these words and then work out a plausible interpretation of them?' That's the wrong way to think about what's going on in the story. Indeed, the question, 'How did Max find out what he meant when he said, "Sarah *is* a philosopher"?' is itself confused. Max didn't *find out* what the first half of his statement meant and then, in the second half of the statement, report what he had learned. We might want to call the second part of Max's statement an interpretation of, or gloss on, the first part. But it's not the sort of gloss or interpretation that another person might provide. The second part of Max's statement is as much a *part* of his assessment of Sarah *qua* philosopher as the first part is. We should think of the two parts of Max's statement as constituting a kind of unity, a single act. I'll describe them as *co-expressing* his assessment of Sarah *qua* philosopher. If a listener wants to understand this assessment, she needs to consider the two parts of Max's statement (the part in which he says, 'Sarah *is* a philosopher', and the part in which he glosses what he meant by these words) as making sense together, as two pieces of a single, extended expression of his thought.

Now consider again the story I told about Max and Sarah at the restaurant. Recall that when Sarah asks Max whether he's smiling at her joke, he replies, 'I'm smiling because I see the waiter with our food, and I'm really pleased that it's coming out so quickly'. Max's smile, I said, is a conscious expression of his pleasure at the imminent arrival of dinner. I

Subscriber: Goldsmiths College, University of London; date: 08 March 2019

want to suggest that we understand Max's gloss on what his smile expresses in the same way that we understood Max's gloss on what he meant by the words, 'Sarah *is* a philosopher'. When Max says what his smile expresses, he is not offering the sort of interpretation of his behaviour that another person might offer. Rather, his smile and his gloss on its significance are two parts of a unified expressive act; together, the smile and the gloss co-express his pleasure at seeing dinner on its way. What I'll call a 'co-expressive gloss' on the psychological significance of, e.g. a smile is not an interpretation of the smile based on observation. It is, rather, a kind of extension of the smile into a linguistic register.

***

At this point, I can say what it is that distinguishes conscious expressions from unconscious ones. Someone's expression of a mental state is conscious if she is able to gloss it—i.e. gloss its psychological significance—co-expressively. Someone's expression of a mental state is unconscious if she is unable to gloss it co-expressively. What it means to say that Max's passive-aggressive dinner-party remarks are unconscious expressions of his anger at Sarah is this: they are expressions of anger at Sarah that Max is unable to gloss co-expressively.

It's not impossible that someone should unconsciously express her frustration by slamming a door and, immediately thereafter, accurately interpret the significance of this behaviour in the way that a psychologically astute external observer might. But in such a case, the door-slam and the gloss on it would be separate acts. When someone consciously expresses her frustration by slamming a door, she is able to provide a gloss on the slam, where such a gloss would, itself, be a continuation—a kind of extension or stretching out—of her expressive act.

7. Replacement

Earlier in this chapter (§2), I discussed §244 of *Philosophical Investigations*, in which Wittgenstein introduces the idea that when a child learns to talk about his pains, he thereby learns a new way to express them. I noted that §244 concludes in a way that's liable to sound off-key, with Wittgenstein saying to his interlocutor, '[T]he verbal expression of pain replaces crying'. I went on to make a suggestion that I was not yet in a position to spell out, viz. that even if 'replaces' ('ersetzt') seems the wrong word for Wittgenstein to have used in this context, there *is* a sense in which the kind of crying that children do as prelinguistic infants is—when they learn to talk—replaced by another kind of crying. I want to return to that suggestion and try to explain it in light of what I just said about the distinction between conscious and unconscious expressions along with something I said a little while ago, in §4.

In §4, I noted that the distinction between, e.g. anger that can be expressed via self-ascription and anger that cannot be so expressed does not, as it were, get a grip when one is considering the mind of a non-linguistic or prelinguistic animal. I went on to say

that when a child gains the ability to self-ascribe his attitudes and emotions, their form changes. Until then, it doesn't make sense to describe them as conscious or as unconscious. From then on, they are either conscious or unconscious or in between conscious and unconscious. Now, something similar happens when a child gains the ability to co-expressively gloss the psychological significance of his expressions: the form of his *expression* changes. Until then, it doesn't make sense to describe his expressions (of sadness, pleasure, love, desire, anger, etc.) as either conscious or as unconscious. From then on, his expressions are either conscious or unconscious or in between. When Kopi expresses a desire to play with me by dropping a Frisbee at my feet, it is no formal imperfection in her expression that she is unable to extend it into a gloss on itself. But if Sarah discovers Max curled up on their bed, sobbing, and if—when she asks him why he is crying—he is completely unable to say, then his crying could be characterized as imperfectly instantiating the form of human expression. Thus, when children learn to talk, one kind of crying does, after all, replace another.

Perhaps this a good place to point out that in this chapter, I am discussing four distinct kinds of *replacement*: (i) the replacement of prelinguistic desires, fears, etc. with the sort of desires, fears, etc. that are characteristic of linguistic creatures, (ii) the replacement of prelinguistic expressions of desire, fear, etc. with the sort of expressions of desire, fear, etc. that are characteristic of linguistic creatures, (iii) the replacement of unconscious desires, fears, etc. with conscious ones, and (iv) the replacement of unconscious expressions of desire, fear, etc. with conscious ones. A case of type (i) or (ii) does not constitute an example of 'making the unconscious conscious'.[12] Cases of types (iii) and (iv) *are* examples of 'making the unconscious conscious'. A central aim of this chapter is to elucidate the logical relations between cases of types (iii) and (iv)—between, that is, two different ways in which the unconscious may be made conscious.

8. Another Answer to Q

In the story that I told about Max and Sarah's dinner party, Max expresses his anger at Sarah via passive-aggressive remarks. As I noted in §6, that story illustrates that it is possible for someone to unconsciously express his conscious anger. Indeed, there is nothing particularly extraordinary about a person's unconsciously expressing either an unconscious emotion or a conscious one. By contrast, we don't find cases that we'd be inclined to describe as ones in which someone *consciously* expresses an *unconscious* state of mind. This is what we should expect, given what I have claimed about the distinction between conscious and unconscious expressions. In order for Max to consciously express, e.g. his anger toward his mother, he would need to produce an expression of anger whose psychological significance he could co-expressively gloss. But now: consider a case in which Max expresses his anger toward his mother by not answering the phone when he sees that she's calling. If Max *could* co-expressively gloss this expression—if he could say, 'I'm not answering the phone because I'm angry at my mother', and, thereby, extend his non-verbal expression into a gloss—his anger would be conscious.

Summing up a bit: we've seen that conscious states of mind may be expressed either consciously or unconsciously. But unconscious states of mind can be expressed only unconsciously. There is, as it were, no grammatical or logical space for an unconscious emotion or attitude to be consciously expressed. A bit of reflection on this point yields the new answer to Q that I promised at the start of this chapter. In §3, I quoted the following statement from *Expression and the Inner*:

> *Someone's mental state is conscious if he has an ability to express it merely by self-ascribing it. If he lacks such an ability with respect to one of his mental states, it is unconscious.*

> (Finkelstein 2003: 120)

Let's call that first answer to Q 'A1'. What follows is a second answer; call it 'A2'.

> *Someone's mental state is conscious if she has an ability to consciously express it. If she lacks this ability with respect to one of her mental states, it is unconscious.*

A2 might sound as if it involves an uncomfortably tight circle, so I should point out that although I am here, as it were, defining a sort of 'consciousness' in terms of a sort of 'consciousness', the latter sort of consciousness has been further defined in terms of co-expressive glossing.[13]

Although A1 and A2 may be thought of as providing two different criteria for a mental state's counting as conscious, if they both are true—as I believe they are—then any mental state that meets either one of these criteria will meet the other one as well. This is to say: (i) if someone is able to express his, e.g. anger merely by self-ascribing it, then he will be able to consciously express it. And (ii) if someone is able to consciously express, e.g. his anger, then he will be able to express it merely by self-ascribing it.[14] In what remains of this chapter, I'll be particularly interested in a point that is implied by (ii), viz. that if someone is *unable* to express his, e.g. anger merely by self-ascribing it, then he will be *unable* to consciously express it.

***

Given the answer to Q that I summarized in §3 and that is encapsulated in A1, making an unconscious emotion or attitude conscious is a matter of acquiring or regaining an ability to express it merely by self-ascribing it. It is such an ability that, I said, Max still lacks even after he's concluded on the basis of behavioural evidence that he is angry at his mother. At the end of §3, I introduced, and then left hanging, a question concerning how this could be the right thing to say about what it takes for an unconscious state of mind to become conscious. Someone might put this question to me as follows: 'As you have characterized him, Max is able to express his unconscious anger at his mother by stranding her at the airport, by misplacing his phone when he knows that she is likely to call, and, presumably, by doing a wide variety of other things that are liable to annoy her. Given the long list of ways in which he can *already* express this anger, why should his

becoming able to express it in *one more way*—by self-ascribing it—be significant? How could the difference between unconscious and conscious anger come down to *that*?'

I am now in a better position than I was in §3 to address the challenge presented by this question. A couple of paragraphs back, I noted that putting A1 and A2 together yields a point that I put as follows: 'if someone is *unable* to express his, e.g. anger merely by self-ascribing it, then he will be *unable* to consciously express it'. I suggest that this point ought to help us understand why it is that the ability to express one's own state of mind in a self-ascription of it should *not* be considered just one expressive ability on a long list of such abilities. Until Max is able to express his anger at his mother merely by self-ascribing it, there is no way for him to consciously express it at all—either by telling his mother that he is angry at her or by refusing to let her stay with him and Sarah when she visits or by choosing not to answer the phone when he sees that she is calling.[15] If Max were to gain the ability to express his anger at his mother merely by self-ascribing it, this would, as it were, open up the possibility of his consciously expressing it in these and myriad other ways. If only for this reason, such a shift would be significant; it would matter to Max's life.

For all that, acquiring an ability to express his anger toward his mother in a self-ascription of it would not guarantee that Max's subsequent expressions of this anger would be conscious rather than unconscious. As we've seen, a conscious state of mind may be expressed *either* consciously or unconsciously. At the end of §5, I said that psychoanalysis aims—not to replace tears with words, but—to replace unconscious expressions with conscious ones. It is perhaps disappointing that acquiring (or regaining) an ability to express an emotion or attitude in a self-ascription of it does not ensure that it won't continue to be expressed unconsciously. Still, the acquisition of such an ability is a necessary step toward replacing unconscious expressions with conscious ones.[16]

References

Finkelstein, D. (1999). 'On the distinction between conscious and unconscious states of mind'. *American Philosophical Quarterly 36* (2) (April): 79–100.

Finkelstein, D. (2003). *Expression and the Inner*. Cambridge, MA: Harvard University Press.

Finkelstein, D. (2010). 'Expression and avowal'. In Kelly Jolley (ed.), *Wittgenstein: Key Concepts* (pp. 185–198). Durham: Acumen Press.

Kluger, J. (2017). *The Animal Mind*. New York: Time Inc. Books.

Lear, J. (2011). *A Case for Irony*. Cambridge, MA: Harvard University Press.

Moran, R. (2011). 'Psychoanalysis and the limits of reflection'. In Jonathan Lear, *A Case for Irony* (pp. 103–114). Cambridge, MA: Harvard University Press.

Wittgenstein, L. (1953). *Philosophical Investigations*, trans. G. E. M. Anscombe. Oxford: Basil Blackwell.

Wittgenstein, L. (1967). *Zettel*, trans. G. E. M. Anscombe, eds. G. E. M. Anscombe and G. H. von Wright. Oxford: Basil Blackwell.

Wittgenstein, L. (1992). *Last Writings on the Philosophy of Psychology Volume 2*, trans. C. G. Luckhardt and M. A. E. Aue, eds. G. H. von Wright and H. Nyman. Oxford: Blackwell.

Notes:

(1.) See, e.g. Finkelstein 1999 and Finkelstein 2003.

(2.) Although Lear is interested in how my answer to Q might bear on our understanding of psychoanalysis, it's worth mentioning that neither Q nor the answer to it that I've given is *about* psychoanalysis as such. (Freud was not the first writer to recognize that we are sometimes moved by unconscious emotions, preferences, and desires. I take it that the title character in Jane Austen's *Emma*, who realizes at the end of the novel that she's been in love with Mr. Knightly for quite a while, presents at least as good an example of someone whose behaviour is motivated by unconscious states of mind as does, say, Freud's Rat Man).

(3.) I believe that many theorists of 'consciousness' fall into confusion by failing to distinguish between 'conscious of x' and 'x's being conscious'. For a discussion of this point, see Finkelstein 2003, ch. 1.

(4.) In her translation of *Philosophical Investigations*, G. E. M. Anscombe sometimes renders 'äußerung' as 'expression' and sometimes as 'manifestation'. Here, I've changed her 'manifestation' to 'expression'.

(5.) It would take me beyond the scope of what I aim to do in this chapter to explain why Wittgenstein thinks this is misguided. Stepping back from Wittgenstein, I would suggest that any such position will suffer from the conflation that I mentioned in n. 3.

(6.) One difference is that winces, moans, and cries (regardless of whether or not they are produced by a human being who can already talk) are not truth-apt, while self-ascriptions are. Wittgenstein has often been read as wanting to deny this difference and claim that psychological self-ascriptions are neither true nor false. This reading is, I think, both uncharitable and unsupported by the relevant texts. (For a discussion of this point, see Finkelstein 2010).

(7.) 'What if someone has a phobia that prevents him from ever speaking about being angry? Couldn't he nevertheless be consciously angry?' As I understand the concept of expression, not every expression must be a *public* manifestation. (It seems to me that on more than one occasion, I have: (i) tried to tiptoe past a sleeping person, (ii) stubbed my toe, and, finally, (iii) expressed the resultant pain in my toe by *thinking*, 'Ow! That really, really hurts!') It's consistent with my understanding of expression to allow that we

sometimes express our states of mind by self-ascribing them in thought, rather than out loud. So yes; someone who has a phobia about *saying* that he's angry can, on my view, nevertheless be consciously angry.

(8.) Imagine someone's replying to what I said in this sentence as follows: 'If a person expresses his, e.g. anger by self-ascribing it, then what his self-ascription "gives voice to" is just that anger and *not*, as you suggest here, a judgment about it'. In order to bracket this concern, my former self might have rewritten this sentence so that it ended as follows: 'but simply the fact that he is sincerely addressing a question about his own state of mind'.

(9.) Wittgenstein 1967: §§422–6.

(10.) I argue that we should think of conscious and unconscious states of mind as lying at two ends of a continuum in Finkelstein 2003: §5.5.

(11.) This isn't to deny that Kopi might be very conscious *of*, e.g. a piece of cheese that's sitting on a plate on a table just above her head. Nor is it to deny that she might bang her head—perhaps in pursuit of that cheese—and momentarily *lose consciousness*. (In a special edition of *Time* called *The Animal Mind*, Steven Pinker is quoted as saying, 'It would be perverse to deny consciousness to mammals' (Kluger 2017). When I say that Kopi's states of mind are neither conscious nor unconscious, I don't take myself to be disagreeing with Pinker's remark).

(12.) I've characterized cases of types (i) and (ii) as involving changes in form—*from* states of mind and expressions that cannot be described as either conscious or unconscious *to* states of mind and expressions that can be so described. Thus, cases of these types are examples—not of 'making the unconscious conscious', but rather—of 'making what is neither conscious nor unconscious either conscious or unconscious'.

(13.) I am not claiming that there won't turn out to be *some* sort of circularity involved in one or both of my characterizations of the distinction between conscious and unconscious states of mind. My aim is not to provide a reductive account of consciousness. I don't mind circles, as long as they aren't so small as to be uninteresting.

(14.) If it still seems unclear that (i) and (ii) both obtain, what follows might make it seem less so. Let's begin with (ii). Imagine that Max slams his office door and that this is (for him, in this situation) a conscious expression of anger that he's feeling toward a colleague named Jeff. What it means for this to be a *conscious* expression of anger is that he can co-expressively gloss its psychological significance; he can respond to a question like, 'Why did you slam the door?' with an answer like, 'I slammed it because I'm angry at Jeff'—and he can do this in such a way that the gloss is a kind of extension of (rather than an observation report about) the initial expressive act (the door-slam). Now, if Max can do that, then he is (*obviously*, I want to say) able to express his anger merely by self-ascribing it, i.e. merely by saying, 'I'm angry at Jeff'. So, if Max's anger counts as conscious according to A2, it will also count as conscious according to A1. Thus, (ii) would

seem to be true. Now let's consider (i). Imagine that Max is able to express his anger at Jeff merely by saying, 'I'm angry at Jeff'. Such a self-ascription would *itself* be a conscious expression. This is to say: Max would have the ability to co-expressively gloss *its* psychological significance. Suppose that Max expresses his anger at Jeff by saying to Sarah, 'I'm angry at Jeff'. Sarah thinks (for reasons with which we needn't concern ourselves) that in so saying, Max might be trying to make some sort of joke. She replies, 'Are you joking? I'm not getting what you mean'. Here, Max would be in a position to reply with a co-expressive gloss on his initial expression, as follows: 'When I said, "I'm angry at Jeff", I wasn't making a joke; I was actually expressing anger that I'm feeling'. So, if Max's anger meets the criterion for consciousness in A1, it will also meet the criterion in A2. Thus, (i) would also seem to be true.

(15.) While his anger toward his mother remains unconscious, Max might unconsciously express it by not answering the phone when he sees that she is calling. But until he can express that anger merely by self-ascribing it, there is no possibility of his *consciously* expressing it by not answering the phone when he sees that she is calling.

(16.) Versions of this chapter were presented at the University of Murcia (Workshop on Contemporary Rationalist and Expressivist Approaches to Self-knowledge, October 2016) and at the University of Chicago (2016 Fall Colloquium, Department of Philosophy, November 2016). I'm grateful to members of both audiences for their questions and comments, especially Matthew Boyle, Jonathan Lear, Ángel García Rodríguez, Ryan Simonelli, and Anubav Vasudevan. In addition, I'm indebted to Agnes Callard, Matthias Haase, and Malte Willer for reading and commenting on drafts.

David H. Finkelstein

Department of Philosophy, The University of Chicago

Oxford Handbooks Online

Abstract and Keywords

This introduction provides an overview of the five chapters in this section, which explores some of the central issues in psychoanalysis as it relates to phenomenology and science. One such issue concerns the scientific status of psychoanalysis (natural or social? bona fide or tendentious?), and more specifically which between the methods of psychoanalysis and its real-life practice may be considered scientific. One of the chapters examines psychoanalysis as a scientific theory of mind, arguing that psychoanalysis fails to test its theories. Another chapter suggests that many of the central tenets of psychoanalytic theory are evidentially supported by recent developments in empirical neuropsychology. Also discussed are the debate between those who view psychoanalysis as science and those who insist that it rather offers a hermeneutic, how psychoanalysis provides a phenomenology in its articulations of unconscious life, and alternative phenomenological schemes for framing the dynamic unconscious.

Keywords: psychoanalysis, phenomenology, science, scientific theory of mind, empirical neuropsychology, unconscious, psychoanalytic theory

Michael Lacewing and Richard G. T. Gipps

PHILOSOPHICAL disputation concerning psychoanalysis often begins—and unfortunately sometimes ends—with the question of whether psychoanalysis counts as scientific psychology. The claim that it does not is often taken to imply 'too bad for psychoanalysis', rather than, say, 'too bad for scientific psychology', or 'so what else is it then?' But the debate, if it is to be philosophically respectable, cannot simply take for granted some or other unexamined conception of what being a science amounts to. And it will only generate confusion, rather than pertinent critique, if it imports standards inappropriate to its object. What it needs are explicit, workable, and apt conceptions of scientificity. What we need to know when engaging with the debate is, for example, whether we are here talking narrowly about natural science or much more permissively about

Wissenschaft—i.e. a systematic body of knowledge—or something in between? If by 'science' we mean *Wissenschaft*, then in that case anthropology, natural history, art history, and psychoanalysis shall—as and when they accrue knowledge rather than indulge speculation—all be counted as sciences. Freud certainly used the term Wissenschaft to describe his endeavours, but he also tended to think that 'the intellect and the mind are objects for scientific research in exactly the same way as any non-human things' (1933: 159).

But what does the 'same way' mean in the quotation from Freud? Natural history differs considerably from, say, chemistry in its use of merely observational rather than experimental methods to accrue knowledge. Freud's suggestion assimilates viable understanding of matters psychological to the model of natural sciences; it leaves out of consideration, for example, the possibility that psychoanalysis may better be considered a social science—a science whose objects and methods are importantly unlike those populating, and used to interrogate, non-human nature. We should also consider that even parents with the best intentions and the most delightful offspring can sometimes impose on their children in such a way as risks thwarting rather than nurturing their flourishing. This, at least, is the force of Habermas's (1987) contention that, in relation to his own intellectual progeny, Freud laboured under 'a scientistic self-misunderstanding'.

(p. 350) The question of the scientific status of psychoanalysis (natural or social? bona fide or tendentious?) also requires specification as to the relevant aspect of its object. Are we talking about the metapsychology (e.g. about putative underlying psychological structures such as the id and superego), the aetiological theories (e.g. as concern the developmental origins of various psychopathologies), the clinical theory (e.g. relating particular difficulties to particular current defensive organizations), or the successfulness of the clinical practice (e.g. as evaluated in outcome studies)? And are we talking about whether the methods of psychoanalysis could, in theory, count as scientific or about whether the real-life practice of psychoanalysts does so? And if we are interested in testing the theory, then what kind of procedure is here to count as 'testing'? And if we are talking about the clinical theory, then which aspect of which one? Psychoanalysis has after all developed a large number of theories of various phenomena from which a variety of hypotheses may be drawn for testing, albeit all organized around centrally recurring themes of dynamically unconscious motivation, defence mechanisms, symbolization, and transference.

This last point raises a deeper question about how to understand psychoanalysis. When we are talking about just such fundamentally organizing conceptions as unconscious motivation, etc. are we describing the a posteriori hypotheses of a fallible science, or rather the a priori organizing conceptualizations which belong to something we could call the 'philosophy' of psychoanalysis? (Compare: do concepts such as 'force' and 'space' function to define physicists' very field of enquiry, or do they function by referring, hopefully successfully, to phenomena that show up within that field?) Perhaps that distinction is itself rather too blunt to be helpful, and we should do better to consider the a priori and the a posteriori as separated by degrees rather than by an immutable chasm.

Phenomenology and Science: Introduction

When we meet with an effective piece of clinical psychoanalytical theory, is its effectiveness to be understood in terms of its ability to bring to life or make visible something which beforehand too often went unnoticed—thus enriching the conceptual frameworks with which we articulate human nature, or is it proposing a new causal explanation for something which was perfectly visible all along? With regards to this issue of visibility, a question not often enough asked, which parallels that of the scientific prowess of psychoanalysis, concerns its phenomenological credentials: what wattage of illumination of human nature has really been achieved by psychoanalysis compared with other sciences and humanities?

In this part, the chapter by Eagle focuses on psychoanalysis as a scientific theory of mind. Taking testability as a measure of scientific standing, he takes psychoanalysis to task for its failure to test its theories. While Freudian theories can be used to generate testable hypotheses, Kleinian and post-Kleinian theories are, he suggests, not so readily testable. While they may play a variety of valuable functions inside and outside the clinic, contributing to a scientific theory of the mind is not amongst them. Furthermore, setting aside the theories, the practice of psychoanalysts does not encourage a scientific approach—an objection pressed most frequently by Frank Cioffi (1999). Particular difficulties noted by Eagle include a failure to address confirmation bias, a dearth of decent published case studies, aetiological theories developed merely on the basis of clinical (p. 351) data, concepts being tacitly redefined to preserve the otherwise implausible theories they articulate, and a culture of hostility to research, guild mentality, and gurus.

While Eagle looks at the extant scientific habits of psychoanalysts, Hopkins—a philosopher—aims to demonstrate the scientific significance of psychoanalysis by newly integrating it with attachment theory and neuroscience, arguing that many of the central tenets of psychoanalytic theory are evidentially supported by recent developments in empirical neuropsychology. Hopkins stresses the consiliences between psychoanalytic understandings of the task of assimilating emotionally overwhelming experience (e.g. in dreaming) to maintain reality contact, the development of symptoms when that breaks down and primary process modes of cognition take over, the significance of our attachment histories for developing the neurological structures required for such emotion processing, Freud's early neurological thought on minimizing 'free energy', and the theoretical neuroscientist Karl Friston's more recent conception of the brain as tasked with free energy minimization by producing such 'predictive models' of the perceptually encountered world as minimizing complexity while maximizing accuracy. We note that other such integrations are possible; those who prefer their theoretical neuroscience in non-computational flavours may consult Allan Schore's (2012) synthesis. The value of Hopkins's chapter for this book, though, is to indicate another way in which psychoanalysis may seek to take its place amongst the sciences. As such it shows philosophy in its synoptic mode (see Gipps and Lacewing, 'Introduction: Know Thyself', this volume)—i.e. in its attempt to bring together disparate fields of knowledge with a philosophical eye to the maximal coherence of the proposed integration. We may note too the other scientific integrations now available and featuring in other sections of this

handbook—with anthropology (Narvaez, this volume), with sociology (Jay, this volume); or elsewhere—for example with cognitive and social psychology (Tesser 1991; Westen 2009).

Lacewing's chapter steps back from the scientific details to review the debate between those who would articulate psychoanalysis as science and those for whom it rather offers a hermeneutic. In fact that very distinction is one which—in the course of a careful unpicking of the assumptions at play in the debate of the last forty years between realist and constructionist strands in American psychoanalysis—is deconstructed in the course of his argument. That debate has, he contends, tended to confound its participants because it too often runs together matters which are more fruitfully kept separate. While it is possible to espouse all of realism (rather than constructionism), to respect (rather than deprecate) metapsychology, to believe in mental determinacy (rather than hold to the indeterminacy of the mental), and to portray psychoanalysis as aiming to provide scientific (as opposed to a different kind) of knowledge, these need not stand or fall together and should not in any case be assimilated. Although he is here not primarily making a case for a particular combination of these themes, one possibility that becomes particularly clear over the course of Lacewing's chapter is that of a realist hermeneutic science which respects mental indeterminacy while deprecating metapsychological speculation.

Matters take a somewhat different tack in the chapters by Fuchs and Gipps. If we understand what it is to be scientific in terms of *testability*, these chapters can be understood as highlighting the importance of primarily *non*-scientific aspects of psychoanalytic knowledge. That is, they are interested instead in our most basic *disclosure or characterization* of unconscious life—psychoanalysis as phenomenology. We do not test the aptness of a fundamental conceptual metaphor, picture, or vocabulary of unconscious life against its object, since that procedure presupposes that we have what here we do not have, namely some other conceptual scheme disclosing that object within which the comparison could be framed. Instead we consider its fruitfulness both in stimulating the development of further illuminations, and in warding off irremediably speculative, esoteric, or otherwise fruitless conceptions of our unconscious life.

Gipps aims to draw out how psychoanalysis provides a phenomenology in its articulations of unconscious life. The first half of his chapter makes a series of distinctions between our acknowledgement of and accounting for phenomena, grammar versus fact, revelation versus representation, and poiesis versus posits, and describes the central aspects of psychoanalytic theory in terms of the former element within each pair. The second half considers how a phenomenological understanding of psychoanalytic theory sheds light on the divergence between the (typical) psychoanalyst's and the (typical) psychologist's understanding of psychoanalytic theory and therapy.

Fuchs canvases several alternative phenomenological schemes for framing the dynamic unconscious, and to replace or at least complement the 'depth' metaphors of traditional psychoanalysis, he develops in detail a conception which puts to use Merleau-Ponty's notion of the lived body and Lewin's notion of a life space. For Fuchs and Merleau-Ponty,

the unconscious is sedimented in such bodily habits as we are blind to precisely because they shape our very experience itself.

Between them, these five chapters offer the basis for future work on the philosophical understanding of psychoanalysis. They analyse and deconstruct both psychoanalytic 'theory' and practice, and provide a range of approaches from which the question 'what is psychoanalysis?' may be answered. They seek to identify and preserve what is most valuable about psychoanalytic thought even while urging its development and integration with other philosophical and empirical approaches to understanding human beings.

References

Cioffi, F. (1999). *Freud and the Question of Pseudo-Science*. Chicago, IL: Open Court.

Freud, S. (1933). 'The question of a Weltanschauung'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 22 (pp. 158–182). London: Hogarth Press.

Habermas, J. (1987). *Knowledge and Human Interests*. Boston, MA: Polity Press.

Schore, A. (2012). *The Science of the Art of Psychotherapy*. New York: W. W. Norton.

Tesser, A. (1991). 'Social vs clinical approaches to self psychology: The self evaluation maintenance model and Kohutian object relations theory'. In R. Curtis (ed.), *The Relational Self: Theoretical Convergences in Psychoanalysis and Social Psychology* (pp. 257–281). New York: Guilford.

Westen, D. (2009). 'The cognitive self and the psychoanalytic self: can we put our selves together?', *Psychological Inquiry 3*: 1–13.

Richard G.T. Gipps

Dr Richard Gipps is a clinical psychologist in private psychotherapy practice in Oxford, UK, and an associate of the Philosophy Faculty at the University of Oxford. He convenes the Philosophy Special Interest Group of the Institute of Psychoanalysis, the Oxford Interdisciplinary Seminars in Psychoanalysis, and the Making the Unconscious Conscious seminar series. His research interests lie in psychoanalysis, psychosis, existential phenomenology, and Wittgenstein.

Michael Lacewing

Dr Michael Lacewing is a former Vice-Principal Academic and Reader in Philosophy at Heythrop College, London, an Honorary Reader in Research Department of Clinical, Educational and Health Psychology at University College, London, and a teacher of philosophy and theology at Christ's Hospital School, Sussex. He has published widely in philosophy of psychoanalysis, metaethics and moral psychology,

alongside writing textbooks for A level philosophy and training in Philosophy for Children (P4C).

Oxford Handbooks Online

Abstract and Keywords

This chapter is primarily concerned with the scientific status of psychoanalysis as a theory of psychological functioning rather than as a form of treatment. It argues that certain characteristics of psychoanalysis, including theoretical formulations that are often obscure and untethered to observation, the proliferation of psychoanalytic 'schools', and habits of mind inculcated in psychoanalytic training and education count against according scientific status to psychoanalysis. The chapter also notes that psychoanalytic formulations have generated a good deal of research on psychological functioning and have yielded important insights into the nature of mind. It suggests that psychoanalytic education and training needs to change in ways that succeed in attenuating loyalty to particular psychoanalytic 'schools' and that encourage openness to empirical evidence and research and the operation of other self-corrective processes.

Keywords: scientific status, psychoanalytic training, clinical data, empirical evidence, theory of mind.

Morris N. Eagle

Introduction

IN addressing the question of the scientific status of psychoanalysis one needs to keep in mind: one, that psychoanalysis is a theory of psychological functioning and development as well as a form of treatment, each aspect generating different issues and questions; and two, that there is not a single psychoanalytic theory, but rather a collection of different psychoanalytic theories and 'schools'. The scientific status of Freudian theory has been debated at great length over a period of many years (e.g. Grunbaum 1984; Hopkins 1988; Lacewing 2018; Rosenblatt 1989; Sachs 1989). Far less discussed are the questions of: one, the scientific status of post-Freudian psychoanalytic theories (however, see Eagle

1986), formulations that have varying degrees of logical relationship to each other; and two, the degree to which self-corrective processes relating evidence to theory are embodied in the attitudes and practices of psychoanalytic practitioners (Cioffi 1970; Eagle 2014).

Although I comment on psychoanalytic psychotherapy process and outcome research, my main focus in this chapter will be on psychoanalysis as a theory of mind and psychological functioning rather than as a form of treatment. I think Freud was right when he stated that the claim of psychoanalysis on posterity will lie in its theory of mind rather than in its form of therapeutic practice.

(p. 354) Psychotherapy and Outcome Research

Like any therapeutic approach, the practice of psychoanalysis and psychoanalytic psychotherapy may have little to do with science, but may be largely a matter of craft and of individual talents and characteristics of the therapist, a conclusion supported by the evidence that individual differences among therapists account for far more of the variance in therapeutic outcome than specific techniques (Lambert and Barley 2001). Scientific-empirical questions enter the picture when one directs systematic inquiry into therapeutic outcome and the processes through which the outcomes occur.

There is a cadre of researchers (generally associated with university departments) who have carried out process and outcome research on psychoanalytic and psychodynamic treatment and have identified some of the factors that mediate therapeutic outcome. This focus on accountability is a great advance over reporting outcomes in the form of clinical anecdotes, self-selected clinical vignettes, testimonials, and 'clinical experience'.[1, 2] Thus, if this chapter dealt with the question of whether psychoanalytically oriented researchers are carrying out outcome and process studies on psychodynamic and psychoanalytic psychotherapy, the answer would be a resoundingly positive one. That is not where the problems related to the scientific status of psychoanalysis lie. The problems lie in the facts that one, to the extent that there is any interest on the part of a large part of the psychoanalytic community in such research, it is mainly limited to the demonstration that psychoanalytic psychotherapy 'works'; two, for the most part, there is little or no feedback from the research findings to psychoanalytic training, education, and clinical practice; and three, there is widespread hostility towards empirical research in the psychoanalytic community. Indeed, the very idea of carrying out research on psychoanalytic treatment has often met with indifference or hostility in the psychoanalytic community.[3] There are good reasons for scepticism regarding the direct value of psychotherapy research for clinical work with individual patients. Insofar as research studies often report group findings of a probabilistic nature, it would be mistaken and foolish to think that one can directly and simply apply such findings to the individual case. (p. 355) However, being aware of research findings can inform the background of one's clinical work and can lead one to question one's cherished and unquestioned assumptions.

Complexities in the Evaluation of the Scientific Status of Psychoanalysis

Let me provide an example. There is a significant body of psychoanalytic thought that views transference interpretations as virtually the sole useful interventions in psychoanalytic treatment (e.g. Gill 1982, 1994). However, there is evidence that transference interpretations can have negative effects on therapeutic outcome (e.g. Ryum et al. 2010; Schut et al. 2005) and that transference interpretations are relatively ineffective with patients who are functioning at a relatively high level of interpersonal relations and are effective mainly with patients who show relatively poor interpersonal relations (Hoglend et al. 2006, 2011; Johansson et al. 2010). The latter finding, in particular, contradicts the received wisdom that transference interpretations are likely to be more effective with higher functioning patients. Surely, any responsible clinician would want to be aware of such findings. Further, such findings are more likely to be useful to therapists than general data on the effect of number of transference interpretations on therapeutic outcome. Moreover, the further investigation of the factors involved in accounting for why patients with relatively poor object relations find transference interpretations more helpful, or which subset of patients within the poor object-relations category find transference interpretations especially helpful, would be even more likely to be useful in clinical practice. The more one includes individual difference variables in one's research, the less sharp the division between nomothetic and ideographic approaches to knowledge, and the greater the direct usefulness of research findings in clinical practice. The general point here is that rather than concluding that research findings are of no relevance to the individual clinician, findings generated by relatively sophisticated research designs can indeed be quite useful, even if their usefulness is limited to leading one to reflect on one's long-standing assumptions in clinical work.

The fact is, however, that most psychoanalytic institutes, particularly free-standing ones, neither require the carrying out of research nor, more noteworthily, prepare their candidates to become critical research consumers. Neither of these activities plays any significant role in psychoanalytic training and education in the majority of psychoanalytic institutes. Furthermore, although not true of all psychoanalysts, as noted, there is a pervasive atmosphere of hostility to research in most psychoanalytic institutes. As described by Kernberg (2012), one of the site visitors' criticisms of the Columbia University Psychoanalytic Institute—an institute more supportive of research—was that an emphasis on research would negatively affect a candidate's clinical training and work.

Kernberg's (2015) critique of the training and education practices of most psychoanalytic institutes echoes Glover's similar disaffection as far back as 1952. He writes:

> It is scarcely to be expected that a student who has spent some years under the artificial and sometimes hothouse conditions of a training analysis and whose professional career depends on overcoming 'resistance' to the satisfaction of his training analyst, can be in a favorable position to defend his scientific integrity (p. 356) against his analyst's theory and practice. And the longer he remains in training analysis, the less likely he is to do so. For according to his analyst the

candidate's objections to interpretations rate as 'resistances.' In short there is a tendency inherent in the training situation to perpetuate error.

(Glover 1952: 403)

Glover (1952) also describes the general state of affairs in the psychoanalytic community as follows:

An analyst, let us say, of established prestige and seniority, produces a paper advancing some new point of view or alleged discovery in the theoretical or clinical field. Given sufficient enthusiasm and persuasiveness, or even just plain dogmatism on the part of the author, the chances are that without any check, this view or alleged discovery will gain currency, will be quoted and re-quoted until it attains the status of an accepted conclusion. Some few observers who have been stimulated by the new idea may test it in their clinical practice. If they can corroborate it they will no doubt report the fact; but if they do not, or if they feel disposed to reject it, this scientific 'negative' is much less likely to be expressed, at any rate in public, and so, failing effective examination, the view is ultimately canonized with the sanctioning phrase 'so and so has shown'. In other words, an *ipse dixit* acquires the validity of an attested conclusion on hearsay evidence only.

(Glover 1952: 403)

Quite disturbingly, there appears to be little or no change in the practices and attitudes of the psychoanalytic community more than sixty years after Glover's comments were made.

The psychoanalytic community's interest in research appears to be directed mainly to findings that suggest the effectiveness of a psychodynamic approach. During the time that I was president of the Division of Psychoanalysis (Division 39) of the American Psychological Association in 1995–6 there was little interest in research among most of its members. A sudden interest in research arose in reaction to a spate of published criticisms launched against the lack of evidence for the effectiveness of psychoanalytic treatment. Some Division officers proposed the formation of a committee whose task it was to amass all the evidence available demonstrating the effectiveness of psychodynamic and psychoanalytic treatment. For the most part, the attitude towards research of individuals making this proposal generally ranged from indifference to hostility; they turned to it only for its demonstration and public relations value. They appeared to want to have it both ways, namely, take the position that psychotherapy research is irrelevant to psychoanalysis and, at the same time, turn to it for its public relations value. This attitude towards research on the part of Division 39 members continues to the present day: there is no longer a research section of the Division, partly due to lack of interest in research papers among its members.

To sum up, as far as a psychoanalytic ethos is concerned, the set of attitudes and habits of mind dominant among psychoanalysts and normative at psychoanalytic institutes is

resistant to systematic inquiry into therapeutic process and outcome. Such inquiry, when it does occur, is limited to a search for confirmation.

(p. 357) Psychoanalysis as Theory of Mind, Personality Development, and Psychopathology

Research on psychoanalytic theories and hypotheses has a long history going at least as far back as a study by Pötzl (1917) on the influence of 'indifferent impressions' on dream content and including, among other work, the Rosenzweig (1934) study on repression, the Sears (1943, 1944) studies on Freudian concepts at Stanford; the anthropological investigations of Malinowski (1927), and of Whiting (1960) and Whiting and Child (1953) at Yale; studies on defence and 'repressive style' (e.g. Weinberger 1990; Weinberger et al. 1979); research on ability to delay gratification (e.g. Mischel 1961); the early work on 'experimental neurosis' (e.g. Masserman 1944, 1950), reflecting an attempt to integrate Pavlovian and Freudian ideas (French 1933); research on Winnicott's (1953) concept of the transitional object (e.g. Cohen and Clark 1984); studies on the relationship between the Freudian notion of orality and dependency (e.g. Bornstein, Poynton, and Masling 1985); the Fisher and Greenberg (1996) volume of empirical test of psychoanalytic hypotheses; Masling and colleagues' volumes on empirical test of psychoanalytic theories (Bornstein and Masling 1998; Masling 1983, 1986, 1990; Masling and Bornstein 1993, 1994, 1996); and research on alexithymia (e.g. Taylor and Bagby 2013).

In short, psychoanalytic theory has been remarkably heuristic in generating a large body of empirical research—mainly by non-psychoanalysts outside the context of psychoanalytic training institutes. However: one, much of this research has had to do with Freudian theory, largely due to the susceptibility to empirical testing of at least a wide range of Freudian hypotheses; two, there appears to be a steady decrease outside psychotherapy research in the number of research papers on psychoanalytic concepts and formulations, largely due, I suggest later, to the unsusceptibility of many post-Freudian theoretical formulations to empirical research; and three, the large body of research findings have had little or no impact on psychoanalytic theorizing, perhaps because of the widespread insistence that only clinical data from the psychoanalytic treatment situation are relevant to psychoanalytic theorizing. I turn now to that assumption.

Clinical Data

A virtually canonical assumption in the psychoanalytic literature, cutting across different schools, is that evidence testing psychoanalytic formulations must be generated by the clinical situation. Although in certain contexts, data derived from the clinical situation can be useful in testing certain psychoanalytic hypotheses (see, for example, the research (p. 358) of Luborsky on 'momentary forgetting' (1967, 1973) and the research of Weiss and Sampson et al. (1986) on the relationship between therapist test-passing versus test-

failing and the emergence of warded off contents), as Grunbaum (1984) cogently argued, there are a number of problems with the uncritical use of such data. One problem is illustrated in a paper by Masling and Cohen (1987) in which they follow the course of a number of sessions of psychotherapy carried out by Carl Rogers, who defined his approach as non-directive. In the early sessions the patient's statements seem to be about equally divided between concern with issues of sex and self. Undoubtedly unwittingly, Rogers is silent when the patient talks about sex and tends to respond with a 'Mmm' or tends to look up when the patient talks about issues of self. Over time, the patient's statements about sex decrease and his statements about self increase. This vignette illustrates both the issue of suggestion in the clinical data as well as the problem of confirmation bias, the latter insofar as the patient's behaviour would tend to reinforce Rogers's belief that issues of self are more important than sexual issues for the patient.

Another problem with relying solely on clinical data to evaluate theoretical formulation is that patients' productions in the clinical situation are often ambiguous and subject to different interpretations. Given the analyst's theoretical commitment, he or she is likely to interpret the patient's free associations in accord with that commitment. For example, if one believes that infantile sexual and aggressive wishes culminating in the Oedipus complex are universal and at the core of psychopathology, one will likely find confirming 'evidence' for that belief in the patient's productions (see Peterfreund 1983). And similarly, if one believes, as Kohut (1984) does, that the need for empathic mirroring is universal and that a traumatic lack of empathic mirroring is at the core of psychopathology, one is likely to find confirming 'evidence' for that belief in the patient's productions. Confirmation bias is undoubtedly present in other disciplines. The important question, however, is the degree to which the discipline allows and encourages self-corrective practices (see Lacewing 2013).

An additional problem arises in the use of clinical data to defend etiological theories and hypotheses (e.g. early lack of empathic mirroring leads to lack of self-cohesiveness in adulthood). Such claims obviously require extra-clinical data. The shaky evidential base of psychoanalytic etiological formulations becomes apparent when one considers that virtually all such formulations are based on follow-back retrospective data rather than follow-up prospective data (see Kohlberg et al. 1972). Employing follow-back data, particularly follow-back clinical data, we develop our etiological theories of psychopathology based solely on patients seen in the clinical situation. We do not get data on individuals who may have been subject to the same vicissitudes as the patients we see, but who function quite adequately. This inevitably leads to an overestimation of the general pathogenic significance of the vicissitudes.

For example, in a follow-back study, Robins and McEvoy (1966) found that 59 per cent of adult drug users were truants as juveniles. However, employing follow-up or longitudinal data, they found that, for children who were truants before age fifteen, 26 per cent were drug users as adults. Hence, an etiological theory linking juvenile truancy to adult drug

use that is based solely on follow-back data would be seriously misleading regarding the strength of the association between the two factors.

(p. 359) Or consider this purely hypothetical example: We have data on 1,000 individuals subjected to sexual abuse as children. Ten per cent of the children sexually abused as children are diagnosed with borderline personality disorder (BPD) as young adults. The 100 young BPD adults will be far more likely to be seen clinically than the 900 others. Suppose further that all 100 of these young adults are seen at a particular clinic. The clinicians treating them there may then form an aetiological theory that significantly overestimates the impact of sexual abuse on the development of BPD. Researchers considering the longitudinal data on all 1,000 sexually abused individuals, however, will find that only 10 per cent of sexually abused children develop BPD, and so will conclude that although sexual abuse is, indeed, a risk factor, it must interact with other factors in determining the development of later BPD. Such research then encourages additional research to identify the factors that jointly precipitate BPD. In short, formulations based on clinical data on adult patients rely on follow-back or retrospective data rather than follow-up prospective data and therefore cannot generate adequate developmental theories. Recognizing this limitation and the importance of including methods and data from other disciplines is necessary if one is to develop an adequate psychoanalytic theory (or any other theory) of development, including the development of psychopathology.

Quite remarkably, for the most part, follow-back clinical data from adult patients constitute the primary basis for central psychoanalytic developmental theories, for example, Freud's (1917 [1916–17], 1925 [1924]) positing of a universal Oedipus complex that emerges at a particular age or, as another example, Kohut's (1984) assertion that lack of self-cohesiveness in adulthood is the consequence of early failures to have the infant's and child's needs for empathic mirroring met.

With regard to the former, even if unconscious incestuous wishes were found in all patients, this would hardly constitute adequate evidence for the positing of a universal Oedipus complex characterized by specific wishes and regularly emerging in all children at a particular age. Individuals seen in treatment, perhaps particularly in psychoanalytic treatment, may constitute a selective group different in important respects from other groups. Indeed, there is evidence that intergenerational conflicts (of which the Oedipus complex is an instance) are more likely to be seen in troubled families (Ascherman and Safier 1990; Erikson 1993; Flores, Mattos, and Salzano 1998). In short, an adequate evaluation of the Oedipus complex hypothesis cannot legitimately rest on clinical data, but rather requires longitudinal data with adequate sampling carried out outside the clinical context. Furthermore, the insistence that only clinical data are relevant to testing psychoanalytic hypotheses is but one instance of a stance that obstructs the possibility of refutation.

Not infrequently, one finds the claim in the psychoanalytic literature that clinical data from adult patients are not only an adequate, but the *only legitimate*, data for psychoanalytic developmental formulations. For example, Green (2000), an influential

spokesperson for this position, has distinguished between the 'real child' of empirical observation and the 'true child' of psychoanalysis and written that 'as far as psychoanalysis is concerned, the "real child" is not our concern' (2000: 61); only the 'true child' is relevant to psychoanalysis. What Green means by the 'true child' is 'the reconstructed child of psychoanalysis' (2000: 61). That is, Green seems to be saying that one can infer the (p. 360) course of development and mental states of the child on the basis of the productions of adult patients in psychoanalytic treatment.

If Green is saying that it is the patient's subjective sense of his or her childhood, that is, his or her psychic reality, rather than data from infant observation that are important in clinical work, his argument is well taken. This point, I suspect, would be apparent to anyone engaged in clinical work. Further, this claim does not require a sharp distinction between the 'true' child and the 'real' child. Nor, and this is a critical point, does it suggest that data from infant and child observation are irrelevant to psychoanalytic theories of the nature of psychological development. After all, psychoanalysis is a theory of mind as well as a clinical treatment.

Although Green's view can be interpreted as applying to the context of the clinical situation, the same cannot be said of Winnicott's (1960) statement that 'Indeed, it is not from direct observation of infants so much as it is from the study of the transference in the analytic setting that it is possible *to gain a clear view of what takes place in infancy itself*' (Winnicott 1960: 595, my emphasis). He goes on to say that 'Freud was able to discover infantile sexuality in a new way because he reconstructed it from his analytic work with his psycho-neurotic patients. In extending his work to cover the treatment of the borderline psychotic patient it is possible for us to reconstruct the dynamics of infancy and of infantile dependence, and of the maternal care that meets this dependence' (Winnicott 1960: 595).

While one can perhaps understand the usefulness of this approach in the context of treatment, it seems clear that one cannot legitimately base theories of infant and child development on data derived from the productions of adult patients in treatment—even if these data were not influenced by the therapist's theoretical commitment. If one wants to try to understand infant and child development, one observes infants and children in systematic ways and over long periods of time. It is remarkable that this is even a debated issue in the psychoanalytic community. That it is a debated issue, however, is important information in addressing the role of the habits of mind and attitudes of psychoanalytic theorists in the assessment of the scientific status of psychoanalysis. To sum up, the three problems with the uncritical use of clinical data for general theoretical formulations are: (1) suggestion; (2) confirmation bias; and (3) the inadequacy of follow-back data for etiological formulations.

It is important to note that clinical data can be very useful in formulating and testing a variety of clinical hypotheses and formulations. As noted earlier, this is illustrated in the work of Luborsky (1967, 1973) on 'momentary forgetting' and the research of Weiss and Sampson et al. (1986) on the relationship between therapist 'test-passing' and 'test-

failing' and the emergence of warded off contents (see also Silberschatz 2005). In these studies, the authors rely on a quasi-experimental design that is appropriate to the phenomena being investigated. For example, Luborsky (1967, 1973) found that compared to 'control' segments of clinical sessions in which momentary forgetting was not present, those segments of clinical sessions that preceded momentary forgetting were independently judged to be characterized by conflict and anxiety, suggesting that what one might (p. 361) refer to as 'mini-repressions' occur when conflictual and anxiety-laden material is being dealt with. With regard to the Weiss and Sampson researchers, based on the clinical data, compared to independent judgements of therapist test-failing (i.e. the independent judgement that therapist interventions tended to confirm the patient's unconscious pathogenic beliefs), therapist test-passing (i.e. the judgement that pathogenic beliefs are disconfirmed by the interventions) was reliably associated with the emergence in the clinical sessions of warded-off contents, broadened interests, and reduced anxiety.

Case Studies

Another means by which clinical data can constitute a legitimate means of gaining knowledge and a legitimate method of empirical research is the use of disciplined case studies (Kazdin 1981, 2008). However, as noted by Portuges (1994),[4] the American Psychoanalytic Association Committee headed by Klumpner and Frank (1991) reviewed the sixty most frequently referenced 'psychoanalytic articles' that had been published in three prominent journals between 1869 and 1982 and scrutinized the top fifteen. They reported:

> [n]ot a single one of these 15 papers included any significant amount of ordinary clinical data. Our sample contained no case studies. What evidence was given in support of the conclusions could only be described as sketchy. We found no verbatim examples of and only one dream fragment … Our most frequent descriptions of the 15 papers were, 'No clinical data', 'No evidence', 'Overgeneralized' or 'Assumptions not testable'. In short, it was as if psychoanalytic data were considered so well documented that the reader would be interested only in questions of interpretation.(as cited in Portuges 1994: 12)

There have been attempts to identify and recommend the standards for the case study method that would enhance its reliability and validity as data (e.g. Eagle and Wolitzky 2011; Edelson 1984). However, as is the case with other empirical data, these attempts have had little impact on papers published in psychoanalytic journals. Little has changed since the Klumpner and Frank (1991) report. The clinical data reported in most psychoanalytic papers tend to be limited to self-selected and sketchy clinical vignettes.

(p. 362) Untestability of Formulations

Whatever the limitations of the above kind of theorizing, for the most part the reader can generally understand what formulations are being proposed and perhaps imagine the kind of evidence that might be relevant to assessing these formulations. Unfortunately the psychoanalytic literature is replete with formulations that are so vague and/or so obscure that it is difficult to conceive of any evidence that could play a role in their evaluation. Especially unfortunate is the fact that the authors of these formulations are not obscure figures, but include the most prominent and influential psychoanalytic theorists in the world today.

Consider, for example, the following passage which is representative of Klein's formulations of infant fantasies. Klein (1930: 24–5) writes:

> In my experience sadism reaches its zenith in this [early] phase, which is ushered in by the oral-sadistic desire to devour the mother's breast (or the mother herself) … It is my experience that in the phantasied attack on the mother's body a considerable part is played by the urethral and anal sadism which is very soon added to the oral and muscular sadism. In phantasy the excreta are transformed into dangerous weapons: wetting is regarded as cutting, stabbing, burning, drowning, while the faecal mass is equated with weapons and missiles.

> (Klein 1930: 24–5)

Other than the comment 'It is my experience', Klein offers no evidential basis for the attribution of such florid fantasies to the infant. Indeed, as I have noted, it is difficult to even imagine what kind of clinical evidence could serve as a basis for such attributions. Whereas one can think of evidence that could support Freud's positing of a universal Oedipal stage or Kohut's claim regarding the need for empathic mirroring, it is very difficult to even imagine the kind of evidence that would support Klein's description of the infant's experience. This leaves one with the question of how Klein knows about the experiences of the young infant. It also raises the question of why such theory should be taken seriously, let alone taught at psychoanalytic institutes.

The writings, particularly the late work, of Bion, another highly influential psychoanalytic theorist, fare no better with regard to the issue of testability. In his later writings, Bion became increasingly mystical and preoccupied with the concept of 'O', which he variously defined as the 'Absolute Truth', 'Ultimate Reality', 'infinity', 'noumena or things in themselves', and 'godhead' (Bion 1965: 147–9; 1970: 52). Even those sympathetic to Bion's thought have commented on the problems with his later work. For example, O'Shaughnessy (2005) writes in regard to this work:

> Contradictions have their appeal; breaking the laws of thought and reason bring a quantum of verbal fun. Yet, in scientific writings, such transgressions lead us to anything and everything we fancy—because, as is readily logically demonstrable,

from a contradiction any proposition follows. Texts with contradictions risk an unending proliferation of meanings.

(O'Shaughnessy 2005: 1524–5)

(p. 363) The exegesis of Bion's writings does not help much. For example, in a recent paper sympathetic to Bion's ideas, Rosegrant (2012) refers to the notorious difficulty of Bion's writing style and comments that 'a style that so dominates the message is better taken seriously in itself, not as incidental, but as an integral part of Bion's message' (Rosegrant 2012: 724). He then goes on to describe 'the psychotic mechanisms in Bion's writing' (724). These include 'attacks on linking' and 'denudation of meaning' (725), a major 'massive' example of which is Bion's notorious grid. Rosegrant writes with regard to the grid: 'he may or may not be describing psychologically and clinically useful ideas, but displaying these ideas in a grid is an attack on the usually linked ideas that a grid will simplify and help with understanding' (727).

Rosengrant writes that 'Bion's writing, which on first glance appears to be about psychosis, is on a more fundamental level an inducement to psychosis. Without providing the grounding that would come from alerting the reader that this is about to happen, Bion immerses the reader in the experience he is ostensibly describing: to read Bion is to be psychotic' (Rosegrant 2012: 727). He then goes on to state: 'to recognize that Bion's writing is mad does not mean that it is without value or meaning, but clarifies that its value and meaning are romantic' (727–8). Rosegrant makes clear that what he means by romantic in the present context is essentially a search for the 'mystical one, god, the god head' (729) that Bion refers to as O. Further according to Rosegrant, for Bion, 'the experience of O is the proper goal of analysis' (730).

As other examples of exegeses of Bion's work, consider the following passages from a prominent Kleinian and Bionian psychoanalyst's account of Bion's ideas:

1. 'Bion took the final step in bringing psychoanalysis into alignment with virtually all the fields of knowledge, including mathematics, physics, aesthetics, literature, mythology, philosophy, poetry, religion, archeology, etc.—all for the purpose of putting it on a veritable Internet of a universal epistemology' (Grotstein 1997: 78).
2. 'The major themes that express this profound paradigm shift include the concept and the penumbra of ideas that cluster around "O", the concept of the mystic, the genius, and the "messiah", the invocation of the "religious instinct", the very idea of God and the Godhead, and his explorations of fetal mental life' (Grotstein 1997: 78).
3. 'It is my impression upon reading Bion that this epistemological-ontological-phenomenological cycle takes place as follows: the infantile portion of the personality experiences O, Absolute Truth, as beta-elements, projects them into the analyst who, in containing them, transforms them from O to K (from Truth to Knowledge) and shares this knowledge with the analysand, who, upon this acceptance of the *knowable truth* allows it to become transformed into wisdom, personal O. I term this last act the achievement of the transcendent position' (Grotstein 2004: 1086).

Klein's and Bion's concepts and formulations may be meaningful, evocative, and helpful to many psychoanalytic clinicians. However, that is not the point here. Insofar as this (p. 364) chapter is concerned with the scientific status of psychoanalytic theories, the issues of evidence and empirical testability are, perforce, central. Further, insofar as Klein and Bion are not marginal figures, but rather two of the most prominent and influential theorists in the psychoanalytic world today, with a large group of followers, the fact that many of their formulations are not susceptible to empirical test is a significant factor in evaluating the status of psychoanalytic theory. This judgement, however, does not preclude the value of their work in other contexts.

Progress and Theory Change

From the earliest period of its history to the present day, the emergence of 'isms' and 'schools', including Jungian, Adlerian, Horneyian, Sullivanian, Kleinian, Bionian, Lacanian, ego psychology, modern Freudian, Fairbairnian, and relational psychoanalytic theories, has characterized theoretical developments in psychoanalysis. Furthermore, the emergence of each of these schools is not based on the presentation of evidence that refutes earlier theories or compelling new evidence that supports the new theory. Rather, theory change in psychoanalysis is characterized by the emergence of a charismatic figure who attracts loyal adherents, often culminating in the establishment of psychoanalytic institutes that provide training from a particular theoretical perspective (Eagle 1987, 1995; Eagle and Wolitzky 1989).

Gedo (1984) has noted, accurately, I believe, that 'experienced clinicians will refuse to alter their convictions on the bases of [the researchers'] results. Instead, they would continue to form their psychoanalytic views in direct response to their personal experiences … the belief that psychoanalysis will make progress by validating the best hypotheses through refined scientific methods is implausible. What we need are innovative ideas powerful enough to compel acceptance by significant portions of the analytic community' (Gedo 1984: 514). However liberal one's criteria for scientificity, it is difficult to think of a discipline as scientific where theory changes take place as a function not of systematic evidence but rather of the evocative power and appeal of its formulations and the charisma of the originators of these formulations.

The historical trajectory of psychoanalytic theorizing is not encouraging to anyone who looks for cumulative progress of a scientific sort. Indeed, although in certain areas one may discern progress, many theoretical developments can be characterized as retrogressive in one important sense. That is, whatever the problems and difficulties of Freudian theory, many of its propositions are formulated with sufficient clarity that they are susceptible to empirical test, as evidenced by the fact that much of the empirical research on psychoanalytic propositions is linked to Freudian theory. Contrastingly, as we have seen, Kleinian and neo-Kleinian theories, the most influential current psychoanalytic theories, are formulated in such a way that the possibility of scientific testing does not

arise. In so far as we are concerned with the development of a science of psychoanalysis, this is not an encouraging trajectory.

(p. 365) Falsifiability of Psychoanalytic Propositions

It has been argued that the possibility of falsifiability speaks to the scientificity of a theory, including psychoanalytic theory (Popper 2003/1965; See also Lakatos 1970). For example, arguing against Cioffi's (1970) position that the attitudes and practices of the psychoanalytic community that protect these propositions from falsifiability compromise its scientific status, Farrell (1961, 1984) has proposed that because certain specific psychoanalytic generalizations can be refuted in principle, psychoanalytic theory should be accorded scientific status. There are a number of problems with Farrell's proposal. One problem, already discussed, is that certain core psychoanalytic propositions do not appear to be susceptible to empirical refutation. For example, how would one go about devising procedures that would allow the possibility of refuting Freud's and Klein's claims regarding an inborn death instinct, let alone its influence on psychological development and functioning? The same question can be addressed to the florid fantasies attributed to young infants by Klein or to Bion's claims regarding beta and alpha elements and alpha functioning.

A second problem with Farrell's proposal is that, as Cioffi (1970) has pointed out, 'refutability in principle is not an adequate criterion of the genuinely empirical character of an enterprise' (Cioffi 1970: 472). Were it an adequate criterion for scientific status, astrology, for example, would have to be seen as scientific. For surely, astrological claims are not only refutable, but, indeed, have been refuted. The issue is, as Cioffi (1970) has noted, that astrological 'theorists' not only do not employ procedures that are able to test their claims but, indeed, engage in practices that obstruct refutation. Thus, if being refutable in principle is not accompanied by being refutable in practice, due to a parade of obstructional procedures of the defenders of the claim, refutability in principle becomes an empty criterion. The situation is somewhat analogous to the relationship between the law and the carrying out of the law. Citing Orwell, Remnick (2016) observes that the law itself affords no protection insofar as how laws are implemented depends on the individuals carrying them out and on 'the general temper of the country' (Remnick 2016: 2). We all know of nations which have freedom-affirming constitutions, but dictatorial practices.

Farrell's (1961) specific example of a psychoanalytic proposition that is refutable in principle is the theory of the Oedipus complex. Indeed, not only is there no convincing evidence in support of a universal Oedipus complex, but a good deal of evidence against that claim (see Eagle 2018). However, the refutation of the theory of the Oedipus complex has not resulted in its relinquishment—just as refutations of astrological claims have not resulted in their abandonment.

Complexities in the Evaluation of the Scientific Status of Psychoanalysis

Like Cioffi, Meehl (1993) also cites the practices and habits of mind of members of the psychoanalytic community when he writes: 'One need not conduct a literature search to realize that the divergences in theory and technique among therapists in a broadly (p. 366) "psychoanalytic" tradition are vast, increasing, and show little or no signs of the sort of cumulative self-corrective development characteristic of post-Galilean science' (Meehl 1993: 324). That is, there is no 'self-corrective' feedback from the evidence to the theory. The evidence contradicting the claim of a universal Oedipal stage of development is largely ignored by many Freudian theorists who continue to refer to the Oedipal complex both in their theoretical and clinical papers.

Another reaction to cogent critiques of the Oedipus complex hypothesis has been to retain the concept, but radically alter its meaning so that it no longer bears much resemblance to the original formulation and theoretical claims. Loewald (1979) writes: 'the oedipal attachments, struggles, and conflicts must also be understood as new versions of the basic union-individuation dilemma' (775). In describing 'oedipal attachments' this way, Loewald (1979) essentially redefines the Oedipus complex so that it is understood as entailing conflict between remaining tied to early parental figures versus separation and autonomy (in Mahler's [1968] terms, between symbiosis versus separation-individuation). The existence of unconscious incestuous and death wishes, the defining features of the Oedipus complex, is no longer posited. Rather, reference to such wishes is now understood not as pointing to actual incestuous and death wishes, but rather as *symbolic* expressions of union and separation-autonomy. However, despite gutting its essential features, which include the universality of unconscious incestuous and death wishes, Loewald continues to refer to the Oedipus complex.

As another example of redefinitions of the Oedipus complex, Morehead (1999) writes that although 'human and animal evidence supports the Westermarck hypothesis' (the hypothesis that prolonged propinquity dampens sexual interest), such evidence does not contradict 'recent analytic theory on the Oedipus Complex [which] does not require the existence of a central, powerful, incestuous sexual drive' (347). Like Loewald, Morehead removes a core aspect of the Oedipus complex—the existence of a central powerful sexual drive—but nevertheless insists that it remains a viable theory. In short, both Loewald and Morehead illustrate obstructional attempts to preserve a central psychoanalytic hypothesis through nominal means despite the dearth of confirming evidence and the accumulation of negative evidence.[5]

One could perhaps argue that the issue is largely a verbal one, i.e. that Loewald and Morehead have merely retained the term 'Oedipus complex' to refer to certain dynamics. However, this argument is not convincing. For one thing, Morehead makes clear that his intention is to *preserve* the theory of the Oedipus complex. Perhaps most important, Loewald's and Morehead's manoeuvres obscure the relationship between evidence and theory through nominal means. That is, they obscure the conclusion that there is little or no evidence in support of Freud's positing of universal incestuous and death wishes. Nor is there evidence that how one resolves Oedipal wishes has a decisive influence on psychological development, including development of gender identity and the structure of

the superego. Given this state of affairs, the warranted conclusion would point to the relinquishment of the Freudian theory of the Oedipus complex. Further, (p. 367) there *is* much evidence that issues of union versus separation-individuation and how they are dealt with *do* exert a major influence on psychological development. However, union versus separation-individuation is not a restatement or reconceptualization of the Oedipus complex but, in an important sense, is a *replacement* of Oedipal theory.

Truth, Science, and a Heuristic Discipline

Despite the difficulties identified earlier, it is clear that psychoanalytic theory has had much to offer to those interested in how the mind works. It has attracted the interest of Nobel laureates such as Adrian (1946) and Kandel and gifted scientists such as Gallese, the co-discoverer of mirror neurons. Indeed, both Kandel (2007) and Gallese (personal communication) have reported that at one time in their life, they seriously thought of becoming a psychoanalyst. It is important to add, however, that the inspiration for these scientists was the corpus of Freud's work on the nature of the mind rather than the post-Freudian theorizing I have described.

Despite the problems I have discussed, there have been post-Freudian theoretical developments that have contributed to an understanding of mental functioning. In my view, these contributions include certain aspects of ego psychology (e.g. the importance of certain ego functions such as reflective capacity, ability to delay gratification, and capacity for affect regulation); the importance of what Mahler (1968) refers to as separation-individuation in development; and Bowlby's (1969, 1973, 1980) attachment theory, particularly its account of the basis for the infant–mother bond and the developmental consequences of the particular nature of that bond.

Although much of this work may have been stimulated by psychoanalytic ideas, for the most part research in these areas takes place outside the psychoanalytic context (another reason that psychoanalysts should familiarize themselves with research relevant to psychoanalysis). For example, research and theory on ego functions (although it is not referred to that way in the non-psychoanalytic literature) is primarily cognitive psychology. Similarly, although Bowlby was trained as a psychoanalyst and viewed himself as one throughout his life, virtually all the burgeoning research and theoretical developments in attachment theory have taken place outside psychoanalysis, in domains such as developmental psychology. These examples suggest that many psychoanalytic ideas, including critical reactions to them, have been richly heuristic in stimulating research. That is, to borrow Reichenbach's (1938/1951) admittedly problematic distinction, one can say that many psychoanalytic ideas have been useful in the context of discovery; the problems lie in the context of testing them.

The value of a body of thought is not defined by its scientific status, nor is it the case that only scientific method can generate important insights and truths. I believe that, at its best, a psychoanalytic mode of thought is of great value in a number of ways, and in particular, it may provide important insights into experience and behaviour that are not

(p. 368) readily available through other means. As an example of this phenomenon, I am reminded of a description of an experience provided by Meehl (1984). He describes observing a man and woman leaving a hospital, the woman carrying a Raggedy Ann doll. Meehl becomes unbearably sad and begins sobbing, imagining that the man and woman have just left their very ill child. Upon arriving at his analytic session, still sobbing, his analyst asks Meehl whether he was mean to his daughter before coming to the session, upon which Meehl becomes immediately aware of the reasons for his sadness and sobbing, which then completely cease.

Let me provide another example from my own clinical experience. Many years ago, I was working with a patient, A. G., whose presenting symptom was his obsessive fear that he might be homosexual, which, if true, would represent for him an abominable sin. He was tormented by these thoughts. The symptom appeared after his girlfriend pressed for them to become engaged, with an eye towards marriage. During one session, he reported that after dinner with his mother and father (at the time, he was still living at home), when his mother asked him to put a baking dish back on a high shelf (he was quite tall), his obsessive thoughts about whether he was homosexual got worse. When I asked him whether he had any further thoughts about this incident, he said that putting the baking dish away was something his father used to do. He then remembered that when his mother had recently asked him to mow the lawn—remarking again that this was something his father used to do—the same thing occurred. He then volunteered that after he received a job promotion, his obsessive thoughts reached their peak in frequency and intensity.

Although other interpretations are certainly possible, given other information about A. G.'s experiences and difficulties, a plausible hypothesis close to the clinical evidence is that a central symbolic meaning of all three events—putting the baking dish away, mowing the lawn, and receiving a job promotion—that were associated with an exacerbation of A. G.'s symptoms is that they all involved being placed in an adult role and confronting the challenges and responsibilities that that entails—a role in which he does things his father used to do.

And, indeed, A. G.'s severe anxiety about leaving home, both figuratively and literally, and feeling panicky and engulfed at the thought of marriage, emerged as the central issues in treatment. The point I want to make with this clinical anecdote, similar to Meehl's point about his own experience, is the following one: I know of no theory other than psychoanalytic theory, particularly its formulation of primary process thinking that (in the case of A. G.) enables one to identify the common symbolic meaning among activities as disparate as placing a baking dish on a high shelf, mowing the lawn, and receiving a job promotion. I also do not know of any other theory that provides as clear an insight into the defensive function of a troubling symptom.[6]

(p. 369) Ironically, this very strength of psychoanalytic theory in finding hidden common motives and meanings in disparate activities is also a potential problem. This is so because the meanings identified in the patient's material are often arbitrary, far removed

from the patient's experience, and vary with the analyst's theoretical orientation (see Peterfreund 1983). Therefore, it becomes very difficult to distinguish between arbitrary theoretical impositions on to the clinical material and interpretations warranted by the clinical evidence. As Meehl (1995) puts it, 'the clear message of history is that anecdotal method delivers both wheat and chaff, but it does not enable us to tell which is which' (1019). Being better able to make this distinction would represent a significant contribution. I think the earlier noted work of psychodynamically oriented researchers such as Luborsky and Weiss and Sampson gives us some idea of how such distinctions can be made. However, large sections of the psychoanalytic community need to overcome their indifference or hostility towards these sorts of studies.

Conclusions

1) Although, as Freud and others believed, the development of psychoanalysis as a separate discipline with its own free-standing training institutes and its relative isolation from other disciplines may have made it possible for it to emerge as a coherent body of thought, its status as a separate discipline with its own education and training institutes has constituted a major barrier to theoretical (and perhaps clinical) progress. For the most part, psychoanalytic associations and training institutes have been more a force for the preservation of guild identity and orthodoxy (including pluralistic orthodoxies) than a means for theoretical and clinical progress (Kernberg 2012, 2015).

2) There appears to be a more recent acceptance of empirical research and greater valuing of the importance of establishing links between psychoanalysis and other disciplines in some quarters of the psychoanalytic community. However, this must be broadened beyond a selective search for findings confirming psychoanalytic hypotheses to reflect an equal interest in findings that might *challenge* psychoanalytic formulations and serve as an impetus to theoretical change and progress.

3) In contrast to central aspects of Freudian theory which, at least in principle, lend themselves to empirical test, many post-Freudian theoretical developments, for example those associated with Klein and Bion, are often so arbitrary, obscure, and untethered to observation, that the question of testability becomes moot. Thus, the historical trajectory of psychoanalytic theorizing does not provide a basis for confidence in the prospect of steady and cumulative progress in psychoanalytic theorizing.

4) Theory change in psychoanalysis comes about not by virtue of new findings, but through the emergence of charismatic figures who attract loyal followers who establish psychoanalytic training institutes that promulgate the ideas of new 'schools'.

(p. 370) The general picture that emerges is that of two cultures, one, so to speak, within the psychoanalytic community of psychoanalytic institutes, and the other outside that community, most often associated with university departments. The latter group,

committed to a scientific ethos, carries out research on propositions directly generated or inspired by psychoanalytic theory (e.g. Blatt 2008 on self-definition and relatedness; Cramer 2008, 2012 on defence; Fonagy and Bateman 2006 on reflective capacity and mentalization). The former group, made up of the great majority of psychoanalysts, tends either to have no interest in or to be hostile to research on psychoanalytic formulations. Further, many of the formulations that are most influential in the psychoanalytic community are not susceptible to empirical test. And, perhaps most important, it is this community that shapes the education, training, and therefore, habits of mind of most psychoanalysts.

To sum up, it seems clear that, at its best, a psychoanalytic mode of thought can generate important insights regarding the nature of mind that may not be readily available through other means. However, at its worst, it has generated arbitrary and obfuscatory formulations far removed from observation and unsusceptible to empirical test by any means. Further, given the nature of psychoanalytic training and education, there appears to be no reliable self-correcting mechanism that enables psychoanalysts to distinguish between the plausible and testable, on the one hand, and the obscure and untestable, on the other. Indeed, if one looks at the historical trajectory of psychoanalytic theorizing from Freudian to dominant contemporary theories, it appears that at least as much influential psychoanalytic theorizing is increasingly characterized by obtuse formulation far removed from observation and evidence. Unhappily, this is a grim picture but, I think, an accurate one. I agree with Kernberg (2012, 2015) that necessary modifications in training and education may serve to brighten this picture. I believe that current practices and habits of mind regarding such matters as attitudes towards evidence and testability need to change. The challenge, in my view, is to be able to integrate the best of what a psychoanalytic mode of thought has to offer with the attenuation of loyalties to particular psychoanalytic 'schools' and an openness to empirical evidence and research in the service of self-corrective practices and procedures.

References

Adrian, E. D. (1946). 'The mental and physical origins of behavior'. *The International Journal of Psychoanalysis 27*: 1–6.

Ascherman, L. I. and Safier, E. J. (1990). 'Sibling incest: A consequence of individual and family dysfunction'. *Bulletin of the Menninger Clinic 54*(3): 311.

Berlin, H. A. (2011). 'The neural basis of the dynamic unconscious'. *Neuropsychoanalysis 13*(1): 5–31.

Bion, W. R. (1954). 'Notes on the theory of schizophrenia'. *The International Journal of Psychoanalysis 35*: 13–18.

Bion, W. R. (1965). *Transformations*. New York: Basic Books.

Bion, W. R. (1970). *Attention and Interpretation*. London: Tavistock.

Blatt, S. J. (2008). *Polarities of experience: Relatedness and self-definition in personality development, psychopathology, and the therapeutic Process*. Washington, DC: American Psychological Association Books.

(p. 371) Bornstein, R. F. and Masling, J. (eds) (1998). *Empirical perspectives on the psychoanalytic unconscious: Empirical studies of psychoanalytic theories*, Vol. 7. Washington, DC: American Psychological Association.

Bornstein, R. F., Poynton, F. G., and Masling, J. (1985). 'Orality and depression: An empirical study'. *Psychoanalytic Psychology 2*(3): 241–249.

Bowlby, J. (1969/1982). *Attachment and Loss*, Vol. 1, *Attachment*, 2nd edn. New York: Basic Books. (Original work published 1969).

Bowlby, J. (1973). *Attachment and Loss*, Vol. 2, *Separation*. New York: Basic Books.

Bowlby, J. (1980). *Attachment and Loss*, Vol. 3, *Loss: Sadness and Depression*. New York: Basic Books.

Cioffi, F. (1970). 'Freud and the idea of a pseudo-science'. In F. Cioffi and R. Borger (eds), *Explanation in the behavioral sciences* (pp. 471–499). Cambridge, MA: Cambridge University Press.

Cohen, K. N. and Clark, J. A. (1984). 'Transitional object attachments in early child and personality characteristics in later life'. *Journal of Personality and Social Psychology 46*(1): 106–111.

Cramer, P. (2008). 'Seven pillars of defense mechanisms theory'. *Social and Personality Psychology Compass 2*: 1963–1981.

Cramer, P. (2012). *The development of defense Mechanisms: Theory, Research, and Assessment*. New York: Springer.

Eagle, M. N. (1984). *Recent Developments in Psychoanalysis: A Critical Examination*. New York: McGraw-Hill.

Eagle, M. N. (1986). 'Critical notice: A. Grunbaum's "The Foundations of psychoanalysis" '. *Philosophy of Science 53*(1): 65–88.

Eagle, M. (1987). 'Theoretical and clinical shifts in psychoanalysis'. *American Journal of Orthopsychiatry* 57(2): 175–185.

Eagle, M. (1995). 'The dynamics of theory change in psychoanalysis'. In J. Earman, A. I. Janis, G. J. Massey, and N. Rescher (eds), *Philosophical Problems of the Internal and External World: Essays on the Philosophy of Adolf Grunbaum* (pp. 373–408). Pittsburg, PA: University of Pittsburg Press.

Eagle, M. N. (2011). 'Psychological processes and neural correlates'. *Neuropsychoanalysis 13*(1): 37–42.

Eagle, M. N. (2014). *The Scientific Status of Psychoanalysis*. Paper given at Conference on Psychoanalysis, Philosophy, Religion, and Art. London, October 2014.

Eagle, M. N. (2018a). *Core Concepts in Classical Psychoanalysis: Clinical, Research Evidence and Conceptual Critiques.* London and New York: Routledge.

Eagle, M. N. (2018b). *Core Concepts in Contemporary Psychoanalysis: Clinical, Research Evidence and Conceptual Critiques.* London and New York: Routledge.

Eagle, M. N. and Wolitzky, D. L. (1989). 'The idea of progress in psychoanalysis'. *Psychoanalysis and Contemporary Thought 12*(1): 27–72.

Eagle, M. N. and Wolitzky, D. L. (2011). 'Systematic empirical research versus clinical case studies: A valid antagonism? *Journal of the American Psychoanalytic Association 59*(4): 791–818.

Edelson, M. (1984). *Hypothesis and Evidence*. Chicago, IL: University of Chicago Press.

Ellenberger, H. (1970). *The Discovery of the Unconscious: The History and Evolution of Dynamic Psychiatry*. New York: Basic Books.

Erikson, M. (1993). 'Rethinking Oedipus: An evolutionary perspective of incest avoidance'. *American Journal of Psychiatry 150*: 411–416.

Farrell, B. A. (1961). 'Can psychoanalysis be refuted? *Inquiry 4*: 16–36.

Farrell, B. A. (1984). 'The standing of psychoanalysis'. *Nous 18*(3): 534–541.

(p. 372) Fenichel, O. (1945). *The Psychoanalytic Theory of Neurosis*. New York: W. W. Norton.

Firestein, S. (1978). *Termination in Psychoanalysis*. New York: International Universities Press.

Fisher, S. and Greenberg, R. (1996). *Freud Scientifically Reappraised: Testing the Theories and Therapy*. Oxford: John Wiley & Son.

Flores, R. Z., Mattos, L.F., and Salazano, F. M. (1998). 'Incest: Frequency, predisposing factor and effects in a Brazilian population'. *Current Anthropology 39*(4): 554–558.

Fonagy, P. and Bateman, A. (2006). 'Mechanisms of change in mentalization-based treatment of BPD'. *Journal of Clinical Psychology 62*(4): 411–430.

French, T. (1933). 'Interrelations between psychoanalysis and the experimental work of Pavlov'. *American Journal of Psychiatry 89*: 1165–1203.

Freud, S. (1917 [1916–17]). *Introductory Lectures in Psychoanalysis*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 16 (pp. 243–463). London: Hogarth Press.

Freud, S. (1925 [1924]). 'An autobiographical study'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 20 (pp. 1–74). London: Hogarth Press.

Gallese, V. (pers. comm.).

Gedo, J. (1984). 'Letter to the Editor'. *Review of Psychoanalytic Books 41*(4): 511–515.

Gill, M. M. (1982). *Analysis of Transference*, Vol. 1, *Theory and Technique*. New York: International Universities Press.

Gill, M. M. (1994). *Psychoanalysis in Transition*. Hillsdale, NJ: The Analytic Press.

Glover, E. (1952). 'Research methods in psychoanalysis'. *The International Journal of Psychoanalysis 33*: 403–409.

Goldberg, A. (1978). *The Psychology of the Self: A Casebook*. New York: International Universities Press.

Green, A. (2000). 'Science and science fiction in infant research'. In J. Sandler, A. M. Sandler, and R. Davies (eds), *Clinical and Observational Psychoanalytic Research* (pp. 41–73). London: Karnac Books.

Grotstein, J. S. (1997). 'Bion, the pariah of "O" '. *British Journal of Psychotherapy 14*(1): 77–90.

Grotstein, J. S. (2004). 'The seventh servant: The implication of a truth drive in Bion's theory of "O" '. *The International Journal of Psychoanalysis 85*: 1081–1101.

Grunbaum, A. (1984). *The Foundations of Psychoanalysis: A Philosophical Critique*. Berkeley/Los Angeles, CA: University of California Press.

Hoglend, P., Amlo, S., Marble, A., Bogwald, K. P., Sjaastad, M. C., and Heyerdahl, O. (2006). 'Analysis of the patient–therapist relationship in dynamic psychotherapy: An experimental study of transference interpretations'. *American Journal of Psychiatry 163*: 1739–1746.

Hoglend, P., Hersoug, A. G., Bogwald, K.-P., Amlo, S., Marble, A, Sorbye, O., and Crits-Christoph, P. (2011). 'Effects of transference work in the context of therapeutic alliance and quality of object relations'. *Journal of Consulting and Clinical Psychology 79*: 697–706.

Hopkins, J. (1988). 'Epistemology and depth psychology: Critical notes on the "Foundations of Psychoanalysis" '. In R. Clarke and C. Wright (eds), *Mind, Psychoanalysis, and Science* (pp. 33–60). Oxford: Blackwell.

Johansson, P., Høglend, P., Ulberg, R., Amlo, S., Marble, A., Bogwald, K., and Heyerdahl, O. (2010). 'The mediating role of insight for long-term improvements in psychodynamic therapy'. *Journal of Consulting and Clinical Psychology 78*: 438–448.

(p. 373) Kandel, E. (2007). *In Search of Memory: The Emergence of a New Science of Mind*. New York: W. W. Norton.

Kazdin, A. E. (1981). 'Drawing valid inferences from case studies'. *Journal of Consulting and Clinical Psychology 49*(2): 183–192.

Kazdin, A. E. (2008). 'Evidence-based treatment and practice: New opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care'. *American Psychologist 63*(3): 146–159.

Kernberg, O. F. (2012). 'Suicide prevention for psychoanalytic institutes and societies'. *Journal of the American Psychoanalytic Association 60*: 707–719.

Kernberg, O. F. (2015). 'Resistance and progress in developing research framework in psychoanalytic institutes'. *Psychoanalytic Inquiry 358*: 98–114.

Klein, M. (1930). 'The importance of symbol formation in the development of the ego'. *The International Journal of Psychoanalysis 11*: 24–39.

Klumpner, G. H. and Frank, A. (1991). 'On methods of reporting clinical material'. *Journal of the American Psychoanalytic Association 39*: 459–477.

Kohlberg, L., LaCrosse, G., and Ricks, D. (1972). 'The predictability of adult mental health from childhood behavior'. In B. B. Wolman (ed.), *Manual of Child Psychopathology*. New York: McGraw-Hill.

Kohut, H. (1984). *How Does Analysis Cure?* Chicago, IL: University of Chicago Press.

Lacewing, M. (2013). 'The problem of suggestion: an analysis and solution'. *Philosophical Psychology 26*: 718–743.

Lacewing, M. (2018). 'The science of psychoanalysis'. *Philosophy, Psychiatry, and Psychology 25*(2): 95–111.

Lakatos, I. (1970). 'Falsification and the methodology of scientific research programs'. In I. Lakatos and A. Musgrave (eds), *Criticism and the Growth of Knowledge* (pp. 91–95). New York: Cambridge University Press.

Lambert, M. J. and Barley, D. E. (2001). 'Research summary on the therapeutic relationship and psychotherapy outcome'. *Psychotherapy: Theory, Research, Practice, Training 38* (4): 357–361.

Loewald, H. (1979). 'The waning of the Oedipal Complex'. *Journal of the American Psychoanalytic Association 27*: 751–775.

Luborsky, L. (1967). 'Momentary forgetting during psychotherapy and psychoanalysis: A theory and research method'. *Psychological Issues 5*(2–3): 175–217.

Luborsky, L. (1973). 'Forgetting and remembering (momentary forgetting) during psychotherapy: A new sample'. *Psychological Issues 8*: 29–55.

Mahler, (1968). *On Human Symbiosis and the Vicissitudes of Individuation*, Vol. 1, *Infantile Psychosis*. New York: International Universities Press.

Malinowski, B. (1927). *Sex and Repression in Savage Society*. London: Routledge & Kegan Paul.

Masling, J. (ed.) (1983). *Empirical Studies of Psychoanalytic Theories*, Vol. 1. Hillsdale, NJ: Analytic Press.

Masling, J. (ed.) (1986). *Empirical Studies of Psychoanalytic Theories*, Vol. 2. Hillsdale, NJ: Analytic Press.

Masling, J. (ed.) (1990). *Empirical Studies of Psychoanalytic Theories*, Vol. 3. Hillsdale, NJ: Analytic Press.

Masling, J. M. and Bornstein, F. R. (eds) (1993). *Empirical Studies of Psychoanalytic Theories*, Vol. 4, *Psychoanalytic Perspectives on Psychopathology*.Washington, DC: American Psychological Association.

(p. 374) Masling, J. M. and Bornstein, F. R. (eds) (1994). *Empirical Studies of Psychoanalytic Theories*, Vol. 5, *Empirical Perspectives on Object Relations Theory*. Washington, DC: American Psychological Association.

Masling, J. M. and Bornstein, F. R. (eds) (1996). *Empirical Studies of Psychoanalytic Theories*, Vol. 6, *Psychoanalytic Perspectives on Developmental Psychology*. Washington, DC: American Psychological Association.

Masling, J. and Cohen, I. S. (1987). 'Psychotherapy, clinical evidence, and the self-fulfilling prophecy'. *Psychoanalytic Psychology 4*: 65–80.

Masserman, J. H. (1944). *Behavior and Neurosis*. Chicago, IL: University of Chicago Press.

Masserman, J. H. (1950). 'Experimental neurosis'. *Scientific American 182*: 38–43.

Meehl, P. E. (1973). 'Why I do not attend case conferences'. In *Psychodiagnosis: Selected Papers* (pp. 225–302). Minneapolis, MN: University of Minnesota Press.

Meehl, P. E. (1984). 'Subjectivity in psychoanalytic inference: The nagging persistence of Whilhelm Fleiss Achensee question'. In J. Earman (ed.), *Minnesota Studies in the Philosophy of Science*, Vol. 10, *Testing Scientific Theories* (pp. 349–412). Minneapolis, MN: University of Minnesota Press.

Meehl, P. E. (1993). 'If Freud could define psychoanalysis, why can't ABPP do it? *Psychoanalysis and Contemporary Thought 16*: 299–326.

Meehl, P. E. (1995). 'Commentaries'. *Journal of the American Psychoanalytic Association 43*: 1015–1023.

Mischel, W. (1961). 'Father absence and delay of gratification: Cross-cultural comparison'. *Journal of Abnormal and Social Psychology 63*: 116–124.

Mitchell, S. A. (1998). 'The analyst's knowledge and authority'. *Psychoanalytic Quarterly 67*: 1–31.

Morehead, D. (1999). 'Oedipus, Darwin, and Freud: One big, happy family? *Psychoanalytic Quarterly 68*(3): 347–375.

O'Shaughnessy, E. (2005). 'Whose Bion? *The International Journal of Psychoanalysis 86*: 1523–1528.

Peterfreund, E. (1983). *The Process of Psychoanalytic Therapy: Models and Strategy*. Hillsdale, NJ: Analytic Press (Erlbaum).

Popper, K. R. (2003/1965). *Conjectures and Refutations: The Growth of Scientific Knowledge*. London: Routledge.

Portuges, S. H. (1994). 'Knowledge knots: Problems of psychoanalytic research in clinical psychoanalysis'. Paper presented at the Annual Spring meeting of the Division of Psychoanalysis of the American Psychological Association. Washington, DC, 16 April 1994.

Pötzl, O. (1917). 'The relationship between experimentally induced dream images and indirect vision'. *Psychological Issues 2*: 41–120.

Reichenbach, H. (1938/1951). *Experience and Prediction*. Chicago, IL: University of Chicago Press.

Remnick, D. (2016). 'An American tragedy', **http://www.newyorker.com/news-desk/an-american-tragedy-donald-trump**.

Robins, L. N. and McEvoy, L. (1966). 'Conduct problems as predictors of substance abuse'. In L. Robins and M. Rutter (eds), *Deviant Children Grown Up* (pp. 182–204). Baltimore, MD: Williams and Wilkins.

Rosegrant, J. (2012). 'Why Bion? Why Jung? For that matter, why Freud? *Journal of the American Psychoanalytic Association 60*(4): 721–745.

(p. 375) Rosenblatt, A. (1989). 'Reinspecting the foundations of psychoanalysis: A rejoinder to Adolf Grunbaum'. *Psychoanalysis and Contemporary Thought 12*: 73–96.

Rosenzweig, S. (1934). 'An experimental study of memory in relation to the theory of repression'. *British Journal of Psychology 24*3: 247–265.

Ryum, T., Stiles, T. C., Svartberg, M., and McCullough, L. (2010). 'The role of transference work: The therapeutic alliance and their interaction in reducing interpersonal problems among psychotherapy patients with cluster C personality disorders'. *Psychotherapy: Theory, Research, Practice, Training 47*: 442–453.

Sachs, D. (1989). 'In fairness to Freud: A critical notice of the *Foundations of Psychoanalysis*'. *The Philosophical Review 98*: 349–378.

Schut, A. J., Casonguay, L. G., Flanagan, K. M., Yamasaki, A. S., Barber, J. P, Bedics, J. D., and Smith, T. L. (2005). 'Therapist interpretation, patient–therapist interpersonal process, and outcome in psychodynamic psychotherapy for avoidant personality disorder'. *Psychotherapy: Theory, Research, Practice, Training 42*(4): 494–511.

Sears, R. R. (1943). 'Survey of objective studies of psychoanalytic concepts'. *Social Science Research Council Bulletin* (No. 51). New York: Social Science Research Council.

Sears, R. R. (1944). 'Experimental analysis of psychoanalytic phenomena'. In J. M. Hunt (ed.), *Personality and Other Behavior Disorders* (pp. 306–332). Oxford: Ronald Press.

Silberschatz, G. (2005). 'The control-mastery theory'. In G. Silberschatz (ed.), *Transformative Relationships: The Control-Mastery Theory of Psychotherapy* (pp. 3–30). New York: Routledge.

Spence, D. P. (1994). *The Rhetorical Voice of Psychoanalysis*: *Displacement of Evidence by Theory*. Cambridge, MA: Harvard University Press.

Taylor, E. J. and Bagby, R. M. (2013). 'Psychoanalytic and empirical research: The example of alexithymia'. *Journal of the American Psychoanalytic Association 61*: 99–133.

Weinberger, D. A. (1990). 'The construct validity of the repressive coping style'. In J. L. Singer (ed.), *Repression and Dissociation* (pp. 337–386). Chicago, IL: University of Chicago Press.

Weinberger, D. A., Schwartz, G. E., and Davidson, R. J. (1979). 'Low-anxious, high-anxious, and repressive coping styles: Psychometric patterns and behavioral and physiological responses to stress'. *Journal of Abnormal Psychology 88*: 369–380.

Weiss, J. and Sampson, H. (1986). *The Psychoanalytic Process: Theory, Clinical Observation, and Empirical Research*. New York/London: Guilford Press.

Whiting, J. W. M. (1960). '*Totem and taboo*: A re-evaluation'. In J. Masserman (ed.), *Psychoanalysis and Human Values* (pp. 150–154). New York: Grune and Stratton.

Whiting, J. W. M. and Child, I. L. (1953). *Child Training and Personality: A Cross-Cultural Study*. New Haven, CT: Yale University Press.

Winnicott, D. W. (1953). 'Transitional objects and transitional phenomena.' *The International Journal of Psychoanalysis 34*: 89–97.

Winnicott, D. W. (1958). 'The capacity to be alone'. *International Journal of Psychoanalysis 39*: 416–420.

Winnicott, D. W. (1960). 'The theory of the parent–infant relationship'. *The International Journal of Psychoanalysis 41*: 585–595. (p. 375) (p. 376)

Notes:

(^{1}) As Meehl (1995) has observed, there probably is no presumed therapeutic intervention in the history of humankind, including bloodletting, purging, and insulin coma therapy, that has not been associated with personal testimony and anecdotes regarding its beneficial effects.

(^{2}) I would add that accountability is a moral as well as an empirical issue. That is, the clinician has the moral responsibility to be open to the issue of accountability and to avail him- or herself of whatever evidence is available regarding therapeutic outcome and the processes linked to outcome. This is not to say that there cannot be open and informed debates about such issues as criteria for outcome and quality of evidence, including its ecological validity.

(^{3}) Ironically, despite a widespread hostility towards research, meetings at the American Psychoanalytic Association are labelled 'scientific' meetings.

(^{4}) In addition to providing a poignant account of the doubts and struggles experienced in psychoanalytic training by an open-minded psychoanalyst, Portuges also contributes an excellent discussion of the scientific status of psychoanalysis as well as of the relationship between research and clinical practice.

(^{5}) [Eds: For a discussion of Lacan's redefinition of the Oedipus complex, see Boothby (this volume).]

(^{6}) As one of the editors (RG) notes, one could hypothesize that an exacerbation of A. G.'s symptoms in reaction to 'doing things his father used to do' may reflect an unconscious anxiety-ridden desire and fantasy to replace father. That is a plausible clinical hypothesis not necessarily contradicting, but perhaps complementing, my formulation. However, it is a hypothesis that further clinical evidence would either support or fail to support.

Morris N. Eagle

California Lutheran University, USA

Oxford Handbooks Online

Abstract and Keywords

Core concepts of psychoanalysis admit integration with recent work in neuroscience and developmental psychology. Karl Friston's free energy neuroscience embodies analogues of Freud's primary and secondary processes, and despite conceptual differences assigns the same causal role as Freud to the minimization of free energy. This supports parallel psychoanalytic and neuroscientific accounts of development, dreams, and symptoms, with the latter expressible as a complexity theory of dreaming and mental disorder. Overall the generative model that performs predictive processing discharges the functions—including the regulation of waking and dreaming consciousness—that Freud assigned the ego, while the prototype emotions systems delineated by Jaak Panksepp and other affective neuroscientists discharge those of the drives or id. Together with the emerging role of REM dreaming in memory consolidation, this provides neuroscientific underpinning for psychoanalytic conceptions of the role of memory, emotion, and phantasy in dreams, as illustrated by Freud's dream of Irma's Injection.

Keywords: predictive processing, free energy, ego, drives, complexity, conflict, dreaming, memory consolidation, phantasy

Introduction: Integrating the Sciences of the Mind

Recent developments have made it reasonable to think that some main claims of psychoanalysis can be relatively directly integrated with the framework of contemporary computational and affective neuroscience. Other approaches in psychology, including attachment theory, are also seeking integration with neuroscience; and both neuroscience

and the study of mental disorder are themselves undergoing integration with genetics and Darwinian evolutionary biology. Here we will focus mainly on psychoanalysis and attachment, leaving evolution for consideration elsewhere.[1]

Probabilistic Regulatory Modelling

The integration advanced here is based on the notion of *probabilistic regulatory modelling* associated with the *free energy principle* recently espoused by Karl Friston and his colleagues (see, e.g. Friston 2010).[2] As accords with Conant and Ashby's (1970) proof that a good regulator of any system must model that system, we can understand Darwinian *adaptation* as shaping phenotypic traits of organisms (and therewith the genes that sustain them) to model aspects of the environment so as to interact with them in ways that promote reproductive success. Taken overall these mappings constitute organisms themselves as dynamic models of the environments with which they interact.

In the nineteenth century—and as a scientific version of Kant's claim that consciousness arose from a conceptual synthesis of the manifold of sensory intuition—Helmholtz proposed that the brain naturally conducted such modelling. He argued that the brain accomplished synthesis by embodying a *generative model* of the causes of the data at its sensory receptors—that is, a model of the causes of these sensory data that could predict the impingements and recreate their effects. (Hence, as Helmholtz remarked, objects were 'imagined in the field of vision' as they would have to be 'to produce the same impression on the nervous mechanism' (1865/1924: 2)).

The hypotheses of this model, as in Kant's account of synthesis, included our concepts of the causes and objects of everyday perceptual experience. In Helmholtz's version our conceptual apparatus *unified* and *predicted* the statistical patterns of the data impinging at the receptors, and so explained these data *for the subject*, by *re-representing* the patterns as *conscious perceptual experiences of their causes*.[3] In this Helmholtz used Kant's philosophy to provide an outline account of the working of the brain in generating consciousness, and one which entailed that the basic sensory data were *not* conscious experiences themselves, but rather the contacts at the neural receptors whose patterns the experiences serve to represent and predict.[4]

Freud's Free Energy Neuroscience

Freud framed many of the basic claims of psychoanalysis in the autumn of 1895, shortly after he discovered the use of free association in analysing his own and his patients' dreams. After this both he and his patients sought self-understanding by following and

describing the rapidly changing contents of their own conscious states, in as much detail as possible and without omission or censorship.

This proved a uniquely informative mode of self-disclosure, which even now remains without parallel in any other discipline. An individual's associations made it possible to consider a whole range of what would emerge as motivated behaviours (verbal utterances, non-verbal actions, thoughts, dreams, and symptoms) together and in relation to one another and to the motives that gave rise to them.[5] Together with his own self-analysis this enabled Freud to learn as much about his patients' states of mind as they were able to put into words, and to extend this by framing and testing hypotheses about what their (and his own) associations indicated they could *not* put into words.

Accordingly he rapidly discovered that 'the pathological mechanisms which are revealed in the most careful analysis in the psychoneuroses bear the greatest similarity to dream-processes' (Freud 1895/1950: 336).[6] He began to shift his investigative and therapeutic focus away from what he and Breuer had taken to be veridical but repressed memory, and towards the fundamental and memory-distorting role of imagination and phantasyy.

Freud's first love in research had been the study of the nervous system, in which he had shown unusual distinction.[7] Hence as his new data prompted new hypotheses, he initially tried to formulate them in neuroscientific terms. Within a few weeks he had framed a prescient account of the brain as operating to minimize a form of *free energy* that was specific to the nervous system.

Freud's idea must have originated as a version of Helmholtz's conception of free energy as *the energy in a system that was available for conversion into work*. All physical engines (animal, vegetable, or mineral) operate in accord with the laws of thermodynamics in converting Helmholtz free energy into work. In organic metabolism this conversion is not straightforward, since it involves concurrent processes that are bidirectional, that is, both energy-storing and energy-releasing. But in applying Helmholtz's notion to the brain and mind, Freud adapted it in ways that we can now see to coincide with the distinct notion of *variational free energy*, as this is used in contemporary neuroscience.[8]

Variational Free Energy (FE)

There is a fundamental conceptual problem in applying Helmholtz's account of modelling to the brain. Predicting sensory impingements and recognizing their causes requires computing, for all relevant causes and impingements, the conditional probabilities of *impingements given causes* that enable the prediction of impingements, and the (inverse) probabilities of *causes given impingements* that enable the recognition of causes. According to the Bayesian mathematics of probability, these calculations are intractable: they have no finitary (closed-form) solution. Thus although Helmholtz had described a

process of modelling that the brain seemed to be performing, the task seemed impossible for the brain to perform.

Geoffrey Hinton and his colleagues solved this problem in a series of papers on the *Helmholtz machine* (Dayan et al. 1995) that marked the transformation of connectionist machine learning into the computational neuroscience of the Bayesian brain. Describing the human sensory system as 'a statistical inference engine whose function is to infer the probable causes of sensory input' (1995: 1), they produced a proof of principle as to how the requisite inferences might be performed. This turned on the notion of *variational free energy* (hereafter FE) that they used in computational modelling.

It could be proved that minimizing the FE of a Bayesian probabilistic model of the causes of some data minimized the divergence in probabilities between those the model assigned to the causes and those of the causes themselves: at zero free energy the probability distributions would be the same. Also, however, the parameters (hypotheses) that determined the FE of a model could be adjusted from within that model (as happens in the brain in perceptual learning) in computationally tractable ways. As Hinton realized, it followed that a brain that embodied such a model as Helmholtz had envisaged could ensure that this model enjoyed maximal evidence-based conformity with the causes it represented by continually adjusting its parameters to minimize FE.

The FE of a model is equal to its *computational complexity* minus its *predictive accuracy*. Both are measured with respect to samples of the data the model adjusts itself to learn to predict, which we can take as the sensory contacts that occur over a single waking period. Accuracy is a measure of predictive success. If we imagine the data of waking impingement as points on a graph, and the FE-minimizing hypotheses generated by the model as curves through those points, accuracy would be maximized by a curve that passed through every data point. Complexity, by contrast, reflects the *changes* imposed on the model in working to increase accuracy, as measured by the number of parameters the model is compelled to *adjust* and the extent of the adjustment required. The brain's model thus minimizes FE by *maximizing accuracy* and *minimizing complexity*, thereby tending to make itself the *simplest best predictor* of the data at the receptors.

Friston extended this framework from the perceptual systems to the overall regulation of bodily change and movement that perception subserves. Thus it appeared that the brain so animated the body as to minimize the variational FE of its regulatory model not just by improving prediction (as in perceptual learning), but also by *framing and satisfying predictive hypotheses as to how optimally to move and act*.[9] Such a minimization of FE constitutes life-sustaining activity that is subserved by the processes that store and release thermodynamic free energy (Seth and Friston 2017). In this sense the minimization of FE takes the foreground in understanding the overall energic working of brain and mind.[10] In consequence, as Helmholtz had claimed, 'every movement we make' can rightly be regarded as an experiment testing the predictive success of brain and body in modelling and navigating their environments.

For both Freud and Friston the free energy to be minimized enters the nervous system via sensory impingement, the main source being the inescapable flow of 'endogenous [*interoceptive*] stimuli' that reflect 'the peremptory demands of the internal needs' (Freud 1895/1950: 297). For Friston these stimuli predict departure from the continuously recalculated targets (including homeostatic set points and allostatic adjustments) for FE-minimizing activity (Pellouzio et al. 2015). Hence as Freud speaks of such inputs as creating a 'demand for work' to produce 'specific actions' that 'bind' (inhibit) this free energy, so Friston speaks of 'an imperative to minimize prediction error … through action' (2012: 248) that likewise inhibits the interoceptive sources. And on both accounts such minimization requires the brain to embody a representation or model of the world, including the agent's body (in Freud the 'bodily ego', bounded both internally and externally by the skin), via which the minimizing activities are conducted.

These similarities, as well as others described in what follows, give Freud's conception of free energy the same overall causal profile as Friston's FE. On this basis Carhart-Harris and Friston (2010) sought 'to demonstrate and develop the construct validity of the Freudian concepts', including those of the primary and secondary processes. In what follows, therefore, we will advance their argument by using 'FE' to designate both conceptions of free energy.

FE and Waking and Dreaming Consciousness

Freud hypothesized that both dreams and symptoms functioned to protect the self (or 'ego') from trauma and conflict. They did so, moreover, in the same way: via the fabrication of fictive experiences and beliefs, forms of imagination or phantasy that created an 'alienation from reality' (Freud 1911b: 218ff) that temporarily masked and pacified the conflict or trauma involved. Dreaming, moreover, seemed to present an innate form of consciousness, which might perform a similar function in infancy.

Accordingly Freud hypothesized that the original form of consciousness was *imaginary*, and produced by a *primary process* that was 'in the apparatus from the first' (1900: 603). In infancy this process operated in accord with the *pleasure principle*, buffering the potentially traumatic influx of FE at birth, and paved the way for the *secondary processes*, which operated in accord with the *reality principle*. These enabled the infantile ego to distinguish perceptual experience from wishful imagining, to establish experiential memory, and to learn to perform the 'specific actions' required to inhibit the internal sources of FE. Freud illustrated this by an example in which an infant used the 'wishful cathexis' provided by imagining the breast, to 'experiment' by turning her head so as to secure the nipple (1895: 328–9, 356). This required the infantile ego to moderate (inhibit) the relief provided by phantasy, in order to focus on the *indications of reality* that enabled movement leading to a veridical *experience of satisfaction*.

Thus against the background of the regulation of pleasure and unpleasure by the primary process, the secondary processes introduced the infant—cry by cry, feed by feed, excretion by excretion—to the harsh realities of the world into which she had been born, and also to her own real resources for coping with them. So as the infant gained a greater apprehension of reality (and to the extent that she did so) the secondary processes inhibited and overlaid the primary process in waking life, relegating it to the motor paralysis and sensory attenuation of dreaming.

This benign development, however, remained under threat from frustration, trauma, or conflict. These tended to activate the primary process in waking, and so could create a continued reliance on wishful phantasy that would later show in the alienation from reality (Freud 1911b) characteristic of mental disorder. Hence as Freud later described matters, the phantasies produced by the primary process 'possess *psychical* in contrast to *material* reality; and we gradually learn to understand that *in the world of the neuroses it is psychical reality that is of the decisive kind*' (Freud 1915–7: 368).

The contemporary FE framework provides a remarkably full parallel to these claims. Sleep is a key factor in neurological development, and infants cycle through phases resembling slow wave (SWS) and rapid eye movement (REM) sleep while still in the womb (Kurth et al. 2015). In the last trimester they spend most time in REM, and as Reissland et al. (2011, 2013) observe, their facial expressions indicate both positive (laughter-related) and negative (cry-related) emotions. Accordingly Hobson, Hong, and Friston (2014) take this as a formative period for the brain's generative model. They hypothesize that the brain is 'genetically endowed with an innate virtual reality generator' whose working 'is most clearly revealed in rapid eye movement sleep dreaming'.

We are thus 'born with a virtual reality model' of what we will later discover to be the worldly causes of sensory impingement; and with the onset of sensory activity this virtual model is 'entrained by sensory prediction errors', to become 'a generative or predictive model of the world'. In Friston's account this is accomplished by *active inference*, which effects the perceptual learning, establishing of memory, and inception of FE-minimizing action that Freud assigned to the secondary processes.

This account also parallels post-Freudian (e.g. Kleinian) descriptions of the primary process as *innate expectation-generating phantasy*. Thus Segal (1964) describes the infant as 'approach[ing] reality armed, as it were, with expectations formed by his unconscious phantasy'. These expectations are comparable to scientific hypotheses, in the sense that 'by testing them in reality [the infant] gradually learns which are applicable and which modes of his own functioning enable him to deal with reality'. (Among them are the 'preconceptions' of ideally good and malignantly bad breast/mother/others that the Kleinian tradition regards as the innate basis for thinking about ourselves and others.)

In Friston's terms these expectation-inducing would be the *prior hypotheses* that are altered to accommodate prediction error as the virtual reality generator becomes a 'predictive model of the world'. Alteration of these priors in the first phase of perceptual learning would produce experientially informed *posterior hypotheses* that would then serve as priors for the next phase.

In all this Helmholtz, Segal, Bion, and Friston construe *thinking* in a way that has become familiar in philosophy—as involving tacit or explicit *abduction*, or *inference to the best explanation* (Douven 2011). Helmholtz and Friston take perceptual consciousness to be created by tacit abduction on the part of the brain or generative model; and as conscious agents we continue abductive inference in everyday life and in science. This is the continuity, as Popper remarked, that enables human beings to 'let their theories die in their stead'.[11]

These accounts clearly impose a common framework that links waking, dreaming, development, and mental disorder. On both accounts waking consciousness is underlain by an original imaginary (virtual reality/phantasy) process that later appears mainly in dreaming, and whose operation in waking is constrained by a model or system of representations of the causes of sensory impingement. On both, therefore, a central aspect of development consists in constructing the worldly model whose adherence to reality inhibits and overlays the imaginary process in waking. This, however, entails the possibility originally envisaged by Freud, namely that the development of this realistic orientation may be more or less unsuccessful, leaving the imaginary process to produce the alienation from reality that appears as mental disorder.[12]

Friston's account of FE casts further light on these matters. We noted earlier that the brain's model minimizes FE by maximizing accuracy and minimizing complexity, thereby tending to make itself the simplest best predictor of the data at the receptors. In Friston's account these tasks are roughly divided between waking and sleeping. As a first approximation (omitting the role of imaginative/primary process activity in waking) the brain operates to increase the accuracy of the generative model in waking, and to reduce its complexity in sleep and dreaming.

In this context the most accurate model would be the one best able to predict the currently optimal course of motor response and put it into action. The contrasting importance of complexity emerges if we consider that if an agent's model is too complex —for example if it is working to satisfy imperatives that conflict with one another—such computation may be difficult or impossible. An agent who *needs* a drink, for example, will maximize accuracy by *wanting* and *getting* a drink. But an agent who has numerous conflicting needs may be unable to effect satisfying computation of this kind, and so remain effectively immobile. Hence the requirement, advanced over the lifetime by processing in sleep, of simplifying to predict a single best alternative. As we will see further on, the reduction of complexity in dreaming apparently takes place via the

imaginary attainment of accuracy that Freud described as wish fulfilment (see Pataki, this volume).

We return to dreaming in 'The Complexity Theory of Dreaming and Mental Disorder'. First, however, we must take up the relation of complexity to emotion and conflict.

Emotion Systems

In regulating waking and dreaming consciousness in such a way that experience, thought, and action harmonize to minimize FE, the generative model discharges the tasks that Freud assigned to the ego. In this, however, it is important to see that the brain as modeller or ego is not just a *functional part* of the personal conscious self, but is rather the *overall regulatory governor* of the organisms we are. Among its activities are the production of the conscious experiences in which we regard ourselves as persons who think, feel, choose, and act, as well as the dreams in which we imagine ourselves doing so.

The locus of this ego- or self-creating regulatory activity is the cerebral cortex; and the *subcortical prototype emotion systems*[13] that the cortex develops to regulate over the first year of life play the role that Freud assigned to the drives or id. By contrast with the cortex, the subcortical generators of emotion are fully functional at birth, and their basic architecture in the brainstem is common to all mammals. Watt and Panksepp (2009: 92–3) describe these systems as 'sitting [in the brainstem] over homeostasis proper', where they give rise to attachment, which becomes 'the massive regulatory-linchpin system of the human brain', exercising '*primary* influence' over the prototype systems below.[14]

As described in Alcaro and Panksepp (2011) and Alcaro et al. (2017), these basic systems divide into two overarching and conflicting groups:

1. Those characterized as 'positive/appetitive' or 'Approach/Reward'. As indicated by Panksepp's capitalized functional descriptions, these include SEEKING, PLAY, LUST, and CARE.
2. Those characterized as 'negative/aversive' or 'Avoidance/Conflict'. These include RAGE and FEAR, as well as the attachment-regulating system that Panksepp describes in terms of PANIC, SEPARATION DISTRESS, and GRIEF (hereafter PSG). This latter system generates the punishing internal experiences that ensure the proximity of reproductive cooperators such as parents and offspring.

The systems in each group are co-activating (RAGE and FEAR activate one another, and PSG activates both, and the objects of SEEKING include cooperators in CARE, PLAY, and LUST). By contrast the groups inhibit one another, acting as parts of an overall structure of emotional opponent-processing (Craig 2015); and they generate conflicting dispositions in physiology and behaviour. They are thus sources both of the opposition between

unpleasure and pleasure stressed by Freud, and of the emotional conflicts that he studied and aimed to mitigate—and that we can now see as contribute to insecure (and particularly disorganized) attachment.

These systems coordinate in reducing FE, but via a dialectic that renders them liable to precipitate emotional conflict. This begins with birth, when the negative/aversive systems are activated by the prediction errors that signal impending violations of basic homeostatic expectations previously met from within the mother's body. This challenge to life-sustaining equilibrium is expressed in what Freud regarded as the infant's 'scream'. This communicative action is often the infant's first, and produces a resonating activation in the aversive systems of the mother (Parsons et al. 2014), galvanizing her response. Sensitive response (Howe 2011) coordinates mother and infant in replacing aversive activations with appetitive and rewarding SEEKING, approach, and the consummatory and life-sustaining activity of nursing. This promotes synchronization and bonding between them (Atzila et al. 2017), paving the way for the other approach/reward systems —LUST, CARE, and PLAY—to participate in deepening their emotional relationship.[15]

In this dialectic the infant's and mother's generative models operate in the same way. Equilibrium-threatening increases in FE activate the avoidance/punishment systems, and the model responds with motor activity (an optimal trajectory) that engages the approach/reward systems to minimize it. In infancy this pattern coordinates mother and infant, but the pattern itself is paradigmatic for the operation of the generative model, whether operating realistically or in virtual reality. It develops together with attachment over the first year, as the cortex extends its governance of the emotion systems through phases of axonal growth, synaptic proliferation, and experience-dependent neural pruning. By the beginning of the second year the brain has roughly doubled in size, and reliable patterns in attachment relationships can be discerned in the 'Strange Situation' procedure, to be discussed in the section 'Conflict and Complexity in Disorganized Attachment' further on (see also Howe 2011).

This marks the stabilizing 'regulatory linchpin' of the infant's establishing of 'internal working models' for emotionally significant relationships with the mother and/or other carers. This modelling discharges what Freud described as 'one of the earliest and most important functions of the mental apparatus' namely 'to bind the instinctual impulses which impinge on it, to replace the primary process prevailing in them by the secondary process and convert their freely mobile cathectic energy into a mainly quiescent (tonic) cathexis'' And as Freud also remarks, and we will see later, 'failure to effect this binding would provoke a disturbance analogous to a traumatic neurosis …' (Freud 1920: 33, 34).

Trauma, Conflict, and Complexity

Infancy combines early neurological development with experiential learning, and is therefore uniquely formative for the brain. For this reason early adverse or traumatic experience can have a lifelong impact on mental and physical health (Anda et al. 2006) that is mediated by changes in the working of the brain (McCrory et al. 2017). Early activation of the aversive emotional systems seems a likely cause of what Freud described as 'the first great anxiety-state' associated with the trauma of birth, and later 'the infantile anxiety of longing—the anxiety due to separation from the protecting mother' (Freud 1923: 58).

The early focus of the conflicting emotion systems on the mother accords with psychoanalytic claims that experience in early infancy alternates between prototype-forming representations of very bad and very good objects, whose originals are the parts of the mother's body with which the infant has significant interactions—her breast, her holding presence, her face, eyes, etc. Recent research on memory suggests that the infant's episodic (autobiographical) memory system, centred in the hippocampus and amygdala, is 'highly engaged in the processing of infantile experience and the storing of latent memories' (Alberini and Travaglia 2017; see also Travaglia et al. 2016). These latent memories, moreover, seem comparable to the early 'memories in feeling' described by Melanie Klein (1952: 234). For although they do not admit explicit recall, the emotions linked with them are nonetheless rearoused when similar experiences occur at later times. Despite their latency, such memories can play an important role in experience-driven maturation and development.

This is important for our purposes, for Friston's conception of complexity is internally connected with both *trauma* and *conflict*. Complexity measures *the burden of adjustment* that experience places on the generative model; and since the primary task of the model is to calculate an optimal motor trajectory, as driven by the potentially conflicting prototype systems, this is effectively the *burden of emotional adjustment* (cf. the notion of *affective load* in Levin and Nielsen 2009). Conflict naturally creates complexity, since it involves the activation of *inconsistent* (conflicting) patterns in behaviour, one or both of which perforce require complexity-reducing adjustment. And experiences are traumatic when—as in post-traumatic stress disorder (Enlow et al. 2014) or borderline personality disorder (Mosquera et al. 2014)—they impose requirements for emotional adjustment that the ego/generative model cannot fully meet. The overarching role of the potentially conflicting prototype systems in human motivation thus ensures that for *our* generative models complexity is mainly *emotional* complexity, even when (as, say, in obsessive compulsive disorder or OCD) it shows in ostensibly cognitive forms.

Conflict and Development

We have seen that the early working of the emotion systems can engage mother and infant in a conflictual relationship that predates the infant's understanding. The activation of conflicting emotions becomes emotional conflict *for the infant as subject and agent* with the recognition of self and other as distinct enduring individuals. Experiments suggest that this development—and with it our conception of ourselves as unique persons bounded by physical bodies that move in space and last through time—starts to take hold in the fourth month, as Melanie Klein suggested in her account of the depressive position.[16]

In the prior, pre-objectual phase (Kumin 1996) infants seem to conceive the world as containing a numerically indeterminate multitude of *episodic proto-objects*, comparable to the 'shifting and unsubstantial "tableau"' in terms of which Piaget and Inhelder (1969) described early experience.[17] Since these proto-objects are derived from perception of parts or aspects of the objects perceived, these are also 'part objects' which are anatomically and psychologically incomplete;[18] and since the infant's conception does not yet encompass their actual patterns of spatio-temporal behaviour, they are particularly liable to imaginary transformation in phantasy.

During this phase the infant is also learning the boundaries of the bodily self via sight, touch, and movement, as well as a range of feelings generated via interoception (Fotoupoulou and Tsakaris 2017).[19] So it is *also* indeterminate for the infant whether these episodically encountered proto-objects are mental or physical, or inside or outside the bodily container of the mind or self, as this is bounded by the skin. In consequence the infant can imagine taking the early proto-objects into itself with milk or food; as good or bad causes of feeling inside the self; as expelled from the self, or put into others, with urine, gas, etc. and so on. (Such phantasies populate dreams, and often re-emerge in mental disorder.)

As the generative model extends its spatio-temporal scope the infant increasingly conceives persons as unique and persisting, and accordingly merges proto-objectual representations to represent good and bad aspects of single complex individuals. This requires reorienting the aversive systems from their original focus on the mother, apparently effected by a coordination of identification, projection, and attachment. The realization that others are unique and irreplaceable spurs identification with their valued and caring aspects, and fosters the infant's capacities for guilt and concern, ameliorating the ingroup-facing aspects of the superego. At the same time aspects that cannot be integrated in this way—particularly those derived from early parent–offspring and sibling rivalry—are split off and projected *outside* the circle of familiar kith and kin.

These developments also establish the family as a basic ingroup, preparing the infant for ingroup cooperation for outgroup conflict. Some early signs of this appear at seven to eight months, when infants start to show both angry protest at separation from their

mothers and also fear of strangers (Schaffer 1971). The separation protests show appreciation of the mother as unique and irreplaceable, as well as the refocusing of the emotion systems that this entails. RAGE at the mother now operates in harmony with PSG and FEAR (of strangers) to maintain proximity and attachment within the family. And the projections that originally created the indeterminately numerous bad proto-objects of early infancy are relocated in the alien stranger, and so can now operate to arouse derivatives of RAGE and FEAR in ingroup cooperation for outgroup conflict.[20]

Conflict and Complexity in Disorganized Attachment

We have been taking the early binding (successful inhibition) of free energy hypothesized by Freud to be accomplished by establishing the emotional bonds of attachment. But as noted in the section 'Emotion Systems', Freud also claimed that 'failure to bind' [successfully inhibit] the FE of the drives 'would provoke a disturbance analogous to a traumatic neurosis …' (1920: 64). We can see such a failure, and learn something of its causes, by considering disorganized attachment, as shown in the Strange Situation.

This is a roughly twenty-minute procedure involving eight episodes, numbered (1) to (8); and we can regard the assignment of categories of attachment by this procedure as measuring the functional integration of the rewarding and aversive prototype systems. In (1) the mother (or other carer) and infant enter the room, and in (2) the infant has the opportunity to explore and play with the toys which are there. In (3) a stranger enters the room, and after a short interval attempts to play with the infant. In (4) the mother goes out of the room, leaving the infant with the stranger, who tries, if play has begun, to continue it. Then in (5), the first reunion, the mother returns and the stranger leaves while the mother settles the infant, providing opportunity to return to exploration and play. In (6) the mother again goes out, this time leaving the child entirely alone, and in (7) the stranger enters and again attempts to play with the child. Finally, in (8), the second reunion, the mother again returns to settle the child, who may (or may not) go back to exploration and play.

Thus while episodes (2) and (5) and (8) are such as to activate the rewarding systems that Panksepp identifies as SEEKING and PLAY in the context of maternal CARE, (3) and (4) and (6) and (7) are such as to activate aversive FEAR, PSG, and RAGE, with the latter directed at the caring mother for her part leaving the infant alone and/or at the mercy of the stranger. These conflicting activations, moreover, are such as to produce *regression*, since they repeat those that marked the infant's appreciation of the mother as unique and irreplaceable at seven to eight months. Roughly, infants who are classed as secure at twelve months show less conflict between the rewarding and aversive systems, and

resolve their conflicts more readily, than those who are classed as insecure, that is, as avoidant, ambivalent-resistant, or disorganized.

Infants classed as disorganized show the most striking and apparently irresolvable conflicts. Here are two examples, from a girl and a boy aged eighteen months, as described in Solomon and George (1999: 131).

> In the second reunion Kate approached her mother with her arms outstretched … when she was about two feet away from making contact, she moved her arms to the side and abruptly circled away from her mother like a banking airplane. As she moved away she had a blank, dazed expression on her face.

> In the first reunion, Sam approached his mother with his eyes cast down. When he was about two feet away he looked up at her, rising suddenly and making gasping noises with his breath as he did so. He quickly looked down again, bared his teeth in a half-grimace/half-smile and turned away. Hunching his shoulders and holding his arms and legs stiffly, he tiptoed to the other side of the room. He sat motionless in the chair for 30s, grasping the armrests and staring straight ahead with a dazed expression.

This is an example of the kind of complexity—and the overall trade-off between complexity and accuracy—generated by emotional conflict. The successive stages of the Strange Situation have aroused such disparate and conflicting images of these toddlers' relationships with their mother that their egos/generative models are now unable to determine whether the more accurate response would be to approach or avoid her. On the present account this is because their online working memories are activating disparate and conflicting pre-objectual representations of self and other—representations that their egos/generative models have not yet integrated into single coherent representations of themselves in relations to others as single enduring objects. Hence these representations are overly complex in a literal sense. If we compare the generative model to the guidance system of a self-driving vehicle, then these toddlers would be attempting to compute a route on the basis of disparate and conflicting maps (internal working models) of the territory they are seeking to navigate.

Disorganized infants characteristically achieve a particular form of coherence by three to five years, when they settle into patterns of *coercion* and *control* in relation to their objects. Thus here is Kate at forty-two months playing with a little boy named Trey:

> Pretending to bake a cake in a toy oven, Kate said in a very loud, bossy tone, 'You can't have cake now! Go away: you can't have cake 'till I call you!' She then ordered Trey to get some dishes for her. She frowned when she saw what he brought and scolded him, shaking her finger with her hand on her hips. Trey pretended to eat … Kate yelled 'No! It's not done!' … She ordered him to go away … when he didn't move she pushed him roughly, and he fell to the floor …

(Solomon and George 1999: 150).

And here is Sam at the same age with a friendly little girl called Jenny, who tries to interest him in playing with her doll.

> At first he ignores her overtures but finally he takes the doll she offers … [and] alternates between brushing its hair tenderly and smashing it against the floor. Jenny tries to integrate Sam's behavior into her pretend play … she pretends the doll is saying in a squeaky voice 'Ow, don't do that!' … Jenny tries again, making her doll say 'I'm going on the bus'. Sam grabbed the toy bus before Jenny could put her doll in it, crammed his doll in, and drove the bus away. Jenny gave up trying to play with Sam

> (Solomon and George 1999: 153).

Here the kind of conflict Sam showed towards his mother at eighteen months seems echoed in his attitude towards Jenny's doll, which he alternatively brushes tenderly and smashes against the floor.

Finally we can see what seem to be pre-objective precursors of this kind of complexity in the microanalyses of videotaped face-to-face interactions between mothers and infants described in Beebe and Lachman (2014). These indicate that disorganization in the Strange Situation at twelve-plus months can be predicted from the fourth month—that is, near the inception of the shift to objectual representation the defines the depressive position—by short (2.5 second) episodes in which the mothers involved *fail to recognize or coordinate empathically with their infants' expressions of distress*. Instead the mothers involved respond to their infants' distress with exaggerated surprise or smiling, or again, as distress increases, with facial freezing and/or looking away. Meanwhile the infants themselves—as in the later disorganized patterns discussed previously—respond with *rapidly conflicting expressions of emotion (*e.g. whimpering and smiling within the same second or so); and they grow more distraught as no response is forthcoming.

These are episodes in which the mothers concerned are *not* meeting activations of their infants' aversive systems in the paradigmatic way discussed previously, that is by responses that generate FE-minimizing approach/reward. Rather they seem comparable to what Bion (1962) and Segal (1975) describe as maternal failure to respond in a *containing* way to infants' projections of distress, leaving infants unable to incorporate a capacity for conflict-modifying containment themselves. Emotional misrecognition of this kind can occur in otherwise caring relationships, particularly, it seems, if the carers experienced something comparable in infancy. Still it can prove traumatic, in the sense of contributing to the kind of complexity found in disorganized attachment, and also, say, in borderline personality disorder (BPD), in which histories of trauma and/or disorganized attachment are common (Mosquera et al. 2014).

The Complexity Theory of Dreaming and Mental Disorder

In the *Project* Freud observed that the processes which produced the 'alienation from reality' of mental disorder had 'the greatest similarity' to those that produce dreams. Hobson, Hong, and Friston (2014) argue that the 'virtual reality' dream processes, together with REM (rapid eye movement) sleep and SWS (slow wave sleep), serve to minimize FE by *reducing the complexity* of the generative model. We have seen that this complexity encompasses emotional conflict and trauma; so the virtual reality generator would be doing the same thing—*generating virtual reality to reduce the complexity introduced by trauma and conflict*—by producing the 'alienation from reality' (Freud 1911b: 218ff) that Freud and many others have found characteristic of mental disorder.

We can thus integrate Freud's and Friston's accounts of dreams and symptoms by advancing a *complexity theory* of dreaming and mental disorder. This is the hypothesis that the conflicts and traumas that Freud thought mitigated by recourse to yphantasy/ virtual reality in both dreaming and mental disorder can be seen as forms of the neurocomputational complexity illustrated by disorganized attachment (Hopkins 2016). The claim is thus that mental disorder is the product of such complexity, together with the mechanisms (which apparently include synaptic pruning) which have evolved to reduce it. As well as incorporating Freud's original hypotheses, this account extends them, by relating waking disorders to aberrant complexity reduction in SWS, REM, and dreaming. As described in Hopkins (2016) this is borne out by research on dreaming in depression as well as by work in schizophrenia and a number of other disorders.[21]

How do sleep and dreaming reduce the complexity of the model, and so prepare it to be a more efficient (simpler) predictor of sensory impingement in the next waking period? We can see this by considering the processes of *consolidation* and *reconsolidation* that establish and update the cortically stored *long-term memories* that inform the model. The 'engram complexes' (Tonegawa et al. 2015a, 2015b) that realize memory are networks of synaptically connected neurons. As Freud anticipated, experiential learning systematically alters the patterns and strengths of synaptic connections (together with other aspects of the neurons that regulate their activity), thus determining the nature and contents of the memories they encode. Those that embody a particular memory are distributed over a number of regions in the brain, which may be connected synaptically and/or coordinated by oscillatory patterns (waves) of various kinds (Boccio et al. 2017).

The working of engram complexes involves two roughly dissociable components: the *contextual content* of the memory (who, what, when, where, etc.), as contributed by encoding activity in the hippocampus; and the *emotional significance* or *valence* of the memory, contributed by encoding activity in the amygdala, as related to the hypothalamus and other subcortical structures. The complexity of what is learned is reflected at the synaptic level in individual neurons. Thus Tonegawa et al. (2015a: 94)

report that conditioning, eliminating, and reconditioning fear memories creates opposing patterns of growth and elimination of spines less than 2 millionths of a metre apart on the same dendrites. This suggests that learning creates dynamic opposition, and hence complexity, even at the level of individual synapses—of which there are trillions in the brain.

Accordingly the creation of learning-embodying synaptic connections puts a physiological load on the neurons involved, and one that increases with complexity. Hence what is called *synaptic homeostasis* (Tononi and Cirelli 2014) requires that these connections be *renormalized* (rescaled to levels optimal for the next waking period) in sleep. Friston (2010) takes this as an important part of the reduction of complexity, and there is now strong evidence that SWS (Tononi and Cirelli 2014) and REM (Li et al. 2017) reduce many synaptic connections while maintaining and strengthen others during sleep. According to recent 'active systems' accounts (Dudai et al. 2015; Feld and Born 2017; Rasch and Born 2013) renormalization is implemented in SWS and REM as part of the overall consolidation and reconsolidation of memory in sleep.

In waking the hippocampus acts in concert with the amygdala as 'the search engine of the brain' (Buzsaki 2011), continually constructing the online working memory that enables 'navigation in *cognitive* [as well as physical] *space*'. This includes 'the computation of relationships among perceived, conceived or imagined items', which depends on memory. In addition the hippocampus rapidly encodes the details of waking experience, so that in SWS it can transfer memories from the day to cortical engram complexes for consolidation as long-term memory. This transfer *reactivates* the complexes, rendering them plastic for integration with the new memories being transferred, and rearouses the associated emotions, so that SWS gives way to REM and dreaming.

In dreaming as in waking, the generative model operates to minimize expected FE, and via activation of the SEEKING system. In waking expected FE arises from sensory impingement and the arousals of memory and emotion engaged in increasing accuracy (e.g. by successful satisfaction of need and desire) to reduce it. In dreaming, by contrast, sensory impingement is curtailed, so that the FE sources for complexity reduction are those of memory and emotion in the process of consolidation dreaming.

For this reason dreaming characteristically involves a pattern in virtual reality that mirrors that in waking, but in relation to the arousal of memory rather than sensory experience. The reactivation of emotion-laden (conflictual or traumatic) memories arouses negative/aversive emotions, prompting the *imaginary* prediction/execution of the FE-minimizing appetitive/rewarding trajectories that Freud described as wish fulfilments. Thus for example Freud's dream of Irma's Injection (1900: 106ff),[22] as we will consider in more detail, begins with Irma complaining to Freud of suffering that his psychotherapy has failed to relieve, but ends with diagnostic triumph on his part.

Accordingly REM dreaming is accompanied by the kind of burst-firing of dopamine neurons that marks waking consummatory reward (Dahan et al. 2007). As this occurs in the conditions of maximal neuroplasticity that enable the reactivated old memories to be integrated with the new, the imaginary attainment of reward in dreaming can reshape the emotional valences of the whole engram complex, rendering it less productive of conflict or trauma, and so emotionally optimized to minimize FE by increasing accuracy during the next waking period. Dreaming thus minimizes complexity by an *imaginary* maximization of accuracy during sleep, which optimizes the brain for actual maximization of accuracy in waking.

As would fit with this, Genzel et al. (2015) argue that REM and dreaming function to modify aversive amygdala-related components of memory.[23] This in turn fits with the long-observed positive effects of the restoration of REM dreaming on depression (Cartwright 2010) as well as with recent experiments on engram complexes (Boccio et al. 2017; Tonegawa et al. 2015a, 2015b) that produce dissociations between their amygdala-related emotional valence and their hippocampus-related contextual contents—and including those in which depression-like activity is reduced by manipulations that arouse reward (Ramirez et al. 2015).

Dreaming thus seems the final phase of both the reduction of neurocomputational complexity and the reconsolidation of memory that occur in sleep (and we would expect these processes to be intertwined). On the present account this reduction/reconsolidation yields *a revised integration of emotion and memory, in which the role of emotionally traumatic or conflictual memories has been mitigated by the imaginary experiences of the dream*.

Dreaming Consciousness and the Reduction of Emotional Complexity

This account integrates the reduction of FE (as complexity) in sleep with observations about the large-scale neurological phenomena of SWS, REM, and synaptic homeostasis, downscaling, and pruning. The role it assigns to the generative model—as *reducing the aversive complexity of emotional memory via the imaginary accuracy of the fictive experiences of dreaming*—explicates the significance of dreams in a way that accords with Freud's as well as many other depth-psychological descriptions. For on this account the significance we assign to dreams derives from their role in *unifying the emotional meaning of the past and the present in preparation for the future*, a task they accomplish by *adjusting emotional significances over the fields of memory in which experiential learning is embodied*.

We can readily trace this hypothesized adjustment of emotional significance in the examples in Freud's *Interpretation of Dreams* as well as later psychoanalytic writings. For in Freud's work as well as that of other analysts, the analysis of a dream can be divided into three parts, which are related by association, memory, and emotion. These are:

i. The *manifest content* of the dream, the series of fictive conscious experiences that the dreamer reports as making up the dream. Elements of this are related by free association (and memory and emotion) to

ii. Memories and feelings *from the day before the dream*, whose contents indicate that they played a causal role in the formation of the dream. These in turn are related by memory, emotion, and association, to

iii. Memories from the more remote past, together with deeper and more traumatic and conflicted emotions. The contents of these indicate that they were unconsciously activated together with (ii) the conscious memories and feelings from the day; and these contents can be seen as *transformed* in (i), the manifest content. So Freud describes these as the *latent content* of the dream.

Free association leads from elements of (i) to elements of (ii) and (iii). These connections, and particularly the relations of transformation between (iii) and (i)—what Freud describes as the transformation of latent into manifest content—enable us to gauge the potential impact of the transformations of the manifest dream on the latent memories and emotions that underlie it. We can illustrate this by taking a fragment from Freud's dream of Irma's Injection , and focusing on a single manifest element (which we will follow with the aid of bold type) together with the memories and emotions that Freud associated with it.

(i). From the manifest content of the dream:

… I at once took [Irma] to one side … to reproach her for not having accepted my 'solution' yet … I took her to the window [to examine her] and looked down her throat … I at once called in Dr. M., and he repeated the examination and confirmed it … [saying] 'There is no doubt it is an infection.' … We were directly aware, too, of the origin of her infection. Not long before, when she was feeling unwell, my friend Otto had given her an injection of …trimethylamin (and I saw before me the formula for this printed in heavy type) … **Injections of that sort ought not to be made so thoughtlessly** …And probably the syringe had not been clean.

(Freud 1900, 107)

This element of the manifest content—Freud's claim in the dream that **Injections of that sort ought not to be made so thoughtlessly**—is clearly associated with memories and feelings, both from the dream day, and from the more remote past.

(ii). Memories and feelings from the day before the dream.

As Freud reports, Irma was a family friend and patient he had been treating for hysteria. As the treatment was ending he had 'proposed a solution to [Irma]' about the nature of her illness 'which she was unwilling to accept', and her sessions were broken off at the summer vacation.

On the day of the dream Freud was visited by Otto, the long-standing friend and colleague whom the dream depicts as giving Irma a toxic injection. Otto had been staying with Irma and her family at their country place, where he had been called away to *give someone an injection*, which event evidently reappeared in the dream in an altered form.

> … I asked him how he had found [Irma], and he answered 'She's better, but not quite well' … Otto's words, or the tone in which he spoke them, annoyed me. I fancied I detected a reproof in them … the same evening I wrote out Irma's case history, with the idea of giving it to Dr M … to justify myself … I had been engaged in writing far into the night …
>
> (Freud 1900: 106).

Here a number of connections between the memories and feelings from the day and the content of the dream are clear. The dream depicts Freud as reproving Irma for not having accepted his 'solution', and explains her continued illness as caused by Otto's having given her a toxic injection. Hence in the dream Freud had transformed his own emotionally aversive daytime experience of feeling himself reproved by Otto on behalf of Irma into one of reproving both Irma and Otto. The **Injections of that sort ought not to be made so thoughtlessly** of Freud's dream was the last and most serious of his self-justificatory retaliations for what he had imagined as Otto's reproof on the day of the dream.

(iii). Memories from the more remote past, together with deeper and more traumatic and conflicted emotions. To **Injections of that sort ought not to be made so thoughtlessly** Freud associated as follows (emphasis again supplied):

> … .this sentence in the dream reminded me once more of **my dead friend** who had so hastily resorted to **cocaine injections** … I noticed too that in accusing Otto of **thoughtlessness in handling chemical substances** I was once more touching upon **the story of the unfortunate Mathilde, which gave grounds for the same accusation against myself** …
>
> (Freud 1900: 117).

These associations go directly to past memories freighted with guilt and shame. As Freud says, they indicate that he felt himself guilty of the same kind of thoughtless handling/ injection of toxic substances of which he accused Otto in his dream, with the addition that in his case injections had been lethal. Moreover, he linked this with his own advocacy of

the medical use of cocaine, which (on the basis of his own experience) he had gone so far as to claim was non-addictive. Thus as he says:

> I had been the first to **recommend** the medical use of **cocaine**, in 1885, and this recommendation had **brought serious reproaches down on me**. The misuse of that drug had hastened the **death of a dear friend of mine**

(Freud 1900: 111).

The 'dear friend' was Freud's beloved mentor, the talented scientist Ernst Fleichel von Marxow. He had become addicted to morphine, taken to relieve incurable nerve pain from an amputation consequent on dissecting a cadaver. Freud had recommended that he use cocaine as a non-addictive substitute, and his subsequent addiction, decline, and death was a source of grief and guilt. The 'unfortunate Mathilde' whom Freud associated with him was someone about whom Freud felt a comparable guilt. He had given her injections of sulphanol, considered harmless at the time, and she, too, had died as result.

This death, moreover, was linked with a kind of unconscious emotional conflict on Freud's part that he was to explain much later in his career. Using the name 'Mathilde' as a switch-word, he associated the Mathilde he had killed with his own daughter Mathilde, who had also been seriously ill.

> My patient—who succumbed to the poison—had **the same name** as my eldest daughter. It had never occurred to me before, but it struck me now almost like an **act of retribution on the part of destiny**. It was as though the replacement of one person by another was to be continued in another sense: **this Mathilde for that Mathilde, an eye for an eye and a tooth for a tooth**

(Freud 1900: 111–12).

In considering how his associations had led to these topics, Freud remarked that it seemed he 'had been collecting all the occasions which I could bring up against myself as evidence of lack of medical conscientiousness'. With hindsight we can see that this was the activity of the agentive part of himself that he would later describe as the cruel and moralistic superego, which not only accused him with his most painfully regrettable medical derelictions, but also threatened him with retribution in the form of the death of his own daughter.

This is an example of the turning of aggression against the self to curb aggression within the family that Freud took as one of the main functions of the moralistic superego, and that is particularly evident in the ferocious self-reproaches as well as the disposition to suicide characteristic of introjective depression (Blatt 2004). These memories were thus also a locus of emotional conflict—the conflict between the ego and the moralistic superego in terms of which Freud was later to understand guilt, shame, and suicide.

In psychoanalytic terms we can say that this part of Freud's dream was produced by his ego to relieve guilt and shame generated by his moralistic superego, whose cruelty is shown in the way he feels it to threaten him with the death of his daughter. Similarly in FE terms we can say that the dream was produced by Freud's generative model acting as complexity-reducing virtual reality generator. In this it was operating to reduce the emotional complexity (conflict and trauma, guilt and shame) embodied in Freud's long-term memories, by reconsolidating them under the impact of the imaginary but rewarding and emotionally simplifying experience of innocence and vindication provided by the dream.

Thus in both accounts the mitigation of conflict or reduction of complexity is accomplished by *fictive* experiences, as partly realized by activity of the approach/reward systems, in what we can regard as *instances of self-induced complexity-reducing experiential learning in dreaming consciousness*. And here the agencies involved in the self-induced learning include 'internal objects', or again aspects of the 'internal working models' of self and other, that subserve the working of the ego/generative model.

The FE account thus enables us to describe the role of Freud's memories and associations more fully, and to integrate them with the 'active systems' account of memory consolidation described in the earlier section 'The Complexity Theory of Dreaming and Mental Disorder'. For as already indicated, the memories and feelings in (ii)—those from the dream day—are evidently *those undergoing consolidation in the dream*, that is, those being transferred from the hippocampus to emotionally related engram complexes in the cortex for storage as part of long-term autobiographical memory.

Again, the memories and feelings in (iii) are those that were already resident in the engram complex, whose unconscious activation during the day had set the stage both for Freud's sensitivity to Otto's remark, and for his self-justifying response to it. These are the memories with which the newcomers are undergoing integration, and which are therefore *rearoused and made plastic*, so as to be *reconsolidated in a new integration of memory and feeling*, emotionally simplified—rendered less conflictual and traumatic—by the experience of the dream.

Thus we can construe Freud's analysis as providing an example of the 'active systems' account of memory consolidation and reconsolidation. In this we see three stages:

1. First stage in the consolidation of memory: Freud's memories from the day—the penetration of Otto's remark, Freud's annoyance, his self-justificatory ruminations, and work late into the night—are transferred in SWS from the hippocampus to the cortical engram complexes that gave the events of the day their emotional significance, and with which they are to be integrated as long-term memories.

2. Second stage in the consolidation of memory: This transfer *rearouses* the target complexes, so that they can be *reconsolidated* in integration with the memories of the recent events they influenced. These older and deeper memories—Freud's advocacy of cocaine, his role in the deaths of his patient and friend, etc.—involve guilt, shame, and internal threats of punishment for his role in respect of toxic injections. Under the impact of this dual arousal of memory and (aversive/punishing) emotion, SWS gives way to REM and dreaming.

3. The final stage in the consolidation of memory: Freud's generative model responds to this arousal of aversive memory and emotion by producing the series of imaginary experiences of the dream, culminating in the wish fulfilling experience in which Freud reproaches Otto with **Injections of that sort ought not to be made so thoughtlessly.**

This powerfully rewarding experience of innocence and vindication, intensified by the embodiment of deep symbolic sexual significance,[24] leaves the cortically stored 'what, when, and where' aspects of Freud's memory intact. But it simplifies the complex of which these memories are part by altering their emotional registration in the amygdala. The fictive experiences of the dream thus freed Freud from the depressive and self-justificatory spiral into which he had fallen the day before, providing a kind of *overnight therapy* (Walker and van der Helm 2009), now being elaborated in neuroscientific work on dreaming and depression.[25] As suggested both by traditional psychoanalytic work on symbol formation (Spillius et al. 2011: Pt. 1, sec. 10) and recent work in cognitive neuroscience (Lewis et al. 2018), this can be seen as part of a more pervasive impact of the creative dreaming imagination on waking life and thought.

Freud's dream thus enabled him to address the difficulties he had felt to be accumulating in his practice in a way that was at once more realistic and more creative than his writing out of Irma's case history the evening before. On waking he turned from his sterile preoccupation with trying to justify his past practice to the fruitful one of seeking to understand his role in that practice at a deeper level. This began with his creative activity in analysing his dream, and so engaging consciously with the memories, emotions, and conflicts (particularly with his own superego) that had been roused by Otto's remark. In this he began the deep revisions in his theories and practice that showed weeks later in the *Project*, and that would eventually enable him to address the difficulties of his patients in a more thoughtful way than he had managed with Irma herself—and in particular by taking into fuller account the inner world of their psychic reality, as well as the outer world of their frustration and trauma.

Conclusion

Multiple pathways in contemporary scientific investigation of mind and brain lead to the relations of dreaming, memory, and emotion, and to the transformations that memories undergo in the processing of dreams. This consilience with psychoanalytic findings

indicates the value of collaboration, particularly among psychoanalysis, attachment, and neuroscience. The discoveries about engram complexes discussed in this chapter are already being applied to psychotherapy (Elsey and Kindt 2017). Their application to human dreaming requires further discovery and examination of the emotionally significant memories that are reconsolidated in our dreams.

Freud's method of analysing dreams by systematically comparing their manifest contents with the results of free association makes such discovery and examination possible, and there is no evidence that any other method can yet do so. So progress in this area, for neuroscience as well as the fields of psychology with which it is undergoing integration, may depend on the willingness of non-psychoanalytic researchers to explore the path originally broken by Freud.[26] If others do persevere in this—which is by no means certain, however great the potential benefit—they may come to see that path as leading to a greater understanding of the unconscious working of the FE-minimizing mind and brain.

References

Alberini, C. and Travaglia, A. (2017). 'Infantile Amnesia: A Critical Period of Learning to Learn and to Remember'. *Journal of Neuroscience37*(24): 578–579.

Alcaro, A., Carta, S., and Panksepp, J. (2017). 'The Affective Core of the Self: A Neuro-Archetypical Perspective on the Foundations of Human (and Animal) Subjectivity. *Frontiers in Psychology 8*: 1424. doi:10.3389/fpsyg.2017.01424

Alcaro, A. and Panksepp, J. (2011). 'The SEEKING mind: Primal neuro-affective substrates for appetitive incentive states and their pathological dynamics in addictions and depression'. *Neuroscience and Biobehavioral Reviews 35*: 1805–1820. doi:10.1016/j.neubiorev.2011.03.002

Ali, J., Spence, C., and Bremner, A. (2015). 'Human infants' ability to perceive touch in external space develops postnatally'. *Current Biology 25*: R965–R979.

Anda, R., Felitti, V., Bremner, J., Walker, J., Whitfield C., Perry B., Dube, S., and Giles, W. (2006). 'The enduring effects of abuse and related adverse experiences in childhood. A convergence of evidence from neurobiology and epidemiology'. *European Archives of Psychiatry and Clinical Neuroscience 256*: 174–186.

Atzila, S., Touroutogloua, A., Rudya, T., Salcedoa, S., Feldman, R., Hookera, J., Dickerson, B., Catana C., and Barrett, L. (2017). 'Dopamine in the medial amygdala network mediates human bonding'. *PNAS 114*(9): 2361–2366.

Baillargeon, R., Spelke, E., and Wasserman, S. (1985). 'Object permanence in five-month-old infants'. *Cognition 20*: 191–208.

Baillargeon, R., Stavans, M., Wu, D., Gertner,Y., Setoh, P., Kittredge, A., and Bernard, K. (2012). 'Object Individuation and Physical Reasoning in Infancy: An Integrative Account'. *Language, Learning and Development 8*(1): 4–46. doi:10.1080/15475441.2012.630610

Barrett, L. F. (2017). *How Emotions are Made*. London: MacMillian.

Beebe, B. and Lachman, F. (2014). *The origins of attachment: Infant research and adult treatment*. New York/Hove: Routledge.

Bion, W. (1962). 'A theory of thinking'. *International Journal of Psycho-Analysis 43*: 306–310 (reprinted in *Second Thoughts* (1967), London: Heinemann, pp. 110–119).

Bion, W. R. (1963). *Elements of Psycho-Analysis*. London: Heinemann.

Bivens, L. and Panksepp, J. (2011). *The Archaeology of the Mind*. New York: W. W. Norton.

Boccio, M., Nabavi, S., and Capogna, M. (2017). 'Synaptic Plasticity, Engrams, and Network Oscillations in Amygdala Circuits for Storage and Retrieval of Emotional Memories'. *Neuron 94*: May 17.

Bower, T. (1977). *Development in Infancy*. San Francisco, CA: W. H. Freeman.

Blatt, S. (2004). *Experiences of Depression: Theoretical, Clinical, and Research Perspectives*. Washington, DC: American Psychological Association.

Buzsaki, G. (2011). *Hippocampus. Scholarpedia 6*(1): 1468.

Campos, J., Barret, K., Lamb, M., Goldsmith, H., and Stenberg, C. (1983). 'Socioemotional Development'. In P. Mussen (ed.), *Handbook of Child Psychology*, Vol 3. New York: John Wiley.

Carhart-Harris, R. and Friston, K. (2010). 'The default-mode, ego-functions and free-energy: A neurobiological account of Freudian ideas'. *Brain 133*: 1265–1283.

Cartwright, R. (2010). *The Twenty-Four Hour Mind: The Role of Sleep and Dreaming in Our Mental Lives*. Oxford: Oxford University Press.

Clark, A. (2016). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*, Oxford/ New York: Oxford University Press.

Conant, R. and Ashby, W. R. (1970). 'Every good regulator of a system must be a model of that system'. *International Journal of Systems Science 1*(2): 89–97.

Craig, A. D. (2015). *How Do You Feel: A Moment with your Neurobiological Self*. Princeton, NJ: Princeton University Press.

Craik. K. (1943). *The Nature of Explanation*. Cambridge: Cambridge University Press.

Crespi, B. and Semeniuk, C. (2004). 'Parent-Offspring Conflict in the Evolution of Vertebrate Reproductive Mode'. *The American Naturalist 163*(4).

Dahan, L., Astier, B., Vautrelle, N., Urbain, N., Kocis, B., and Chouvet, G. (2007). 'Prominent burst firing of dopaminergic neurons in the ventral tegmental area during paradoxical sleep'. *Neuropsychopharmacology 32*: 1232–1241. doi:10.1038/sj.npp. 1301251

Damasio, A. and Carvalho, G. (2013). 'The nature of feelings: evolutionary and neurobiological origins'. *Nature Reviews Neuroscience 14*: 143–152. doi:10.1038/nrn3403

Darwin, C. (1871). *The Descent of Man, and Selection in Relation to Sex*. London: John Murray.

Dayan, P., Hinton, G., Neal, R., and Zemel, R.(1995). 'The Helmholtz Machine'. *Neural Computation 7*: 899–904.

Dayan, P. and Abbot, L. (2001). *Theoretical Neuroscience*. London: MIT Press.

Douven, Igor (2011). 'Abduction'. In **The Stanford Encyclopedia of Philosophy** (Spring 2011 edn), E. N. Zalta (ed.).

Dudai, Y., Karni, A., and Born, J. (2015). 'The Consolidation and Transformation of Memory'. *Neuron 88*: October 7.

Elsey, J. and Kindt, M. (2017). 'Tackling maladaptive memories through reconsolidation: From neural to clinical science'. *Neurobiology of Learning and Memory 142*: 108–117.

Enlow, M., Egeland, B., Carlson, E., Blood, E., and Wright, R. (2014). 'Mother–infant attachmentand the intergenerational transmission of post traumatic stress disorder'. *Developmental Psychopathology 26*: 41–65. doi:10.1017/S0954579413000515

Feld, G. and Born, J. (2017). 'Sculpting Memory in Sleep'. *Current Opinion in Neurobiology 44*: 20–27.

Fischmann, T., Russ, M. and Leuzinger-Bohleber, M. (2013). 'Trauma, dream, and psychic change in psychoanalyses: A dialog between psychoanalysis and the neurosciences'. *Frontiers in Human Neuroscience 7*: 877. doi:10.3389/fnhum.2013.00877

Fotopoulou, A. and Tsakiris, M. (2017). 'Mentalizing homeostasis: The social origins of interoceptive inference'. *Neuropsychoanalysis 19*: 3–29. doi: 10.1080/15294145.2017.1294031

Freud, S. (1895/1950). 'Project for a Scientific Psychology'. In J. Strachey (ed.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 1 (pp. 295–324). London: Hogarth Press.

Freud, S. (1900/1953). *The Interpretation of Dreams*. In J. Strachey (ed.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vols 4 and 5. London: Hogarth Press.

Freud, S. (1911b/1957). 'Formulations on the Two Principles of Mental Functioning'. In J. Strachey (ed.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 12 (pp. 1–628). London: Hogarth Press.

Freud, S. (1915–17/1957). 'Introductory Lectures on Psycho-Analysis'. In J. Strachey (ed.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 15 (pp. 218–227). London: Hogarth Press.

Freud, S. (1920/1955). 'Beyond the Pleasure Principle'. In J. Strachey (ed.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 18 (pp. 1–64). London: Hogarth Press.

Freud, S. (1921/1955). 'Group Psychology and the Analysis of the Ego'. In J. Strachey (ed.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 18 (pp. 67–143). London: Hogarth Press.

Freud, S. (1923/1955). 'The Ego and the Id'. In J. Strachey (ed.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 19 (pp. 1–60). London: Hogarth Press.

Friston, K. (2010). 'The free-energy principle:a unified brain theory?'. *Nature Reviews Neuroscience 11*: 127–138. doi:10.1038/nrn2787

Friston, K. (2012). 'Prediction, perception and agency'. *Internatio nal Journal of Psychophysiology 83*: 248–252. doi:10.1016/j.ijpsycho.2011.11.014

Friston, K. (2013). 'Life as we know it'. *Journal of the Royal Society Interface 10*: 0475. doi:10.1098/rsif.2013.0475

Genzel, L., Spoormaker, V., Konrad, M., and Dresler, M. (2015).'The role of rapid eyemovement sleep for amygdala-related memory processing'. *Neurobiology of Learning and Memory 122*: 110–121. doi:10.1016/j.nlm.2015.01.008

Helmholtz, H. (1865/1924). *A Treatise on Physiological Optics*, Vol. III, trans. J. P. C. Southall. New York: Optical Society of America.

Hobson, A., Hong C., and Friston, K. (2014). 'Virtual reality and consciousness inference in dreaming'. *Frontiers in Psychology 5*: 1133. doi:10.3389/fpsyg.2014.01133

Hobson, A. (2014). *Ego Damage and Repair:Towards a Psychodynamic Neurology*. London: Karnac.

Hobson, A. (2015). *Psychodynamic Neurology: Dreams,Consciousness, and Virtual Reality*. New York: Taylor and Francis.

Hohwy, J. (2013). *The Predictive Mind*. Oxford: Oxford University Press.

Hopkins, J. (1987). 'Synthesis in the Imagination: Psychoanalysis, Infantile Experience, and the Concept of an Object'. In J. Russell (ed.), *Philosophical Perspectives on Developmental Psychology*. Oxford: Basil Blackwell.

Hopkins, J. (1999). 'Patterns of interpretation: Speech, action, and dream'. In L. Marcus (ed.), *Cultural Documents: The Interpretation of Dreams* (pp. 123–169). Manchester: Manchester University Press.

Hopkins, J. (2002). 'Psychoanalysis, Metaphor, and the Concept of Mind'. In M. Levine (ed.), *The Analytic Freud* (pp. 22–47). London: Routledge.

Hopkins, J. (2012). 'Psychoanalysis Representation and Neuroscience: The Freudian unconscious and the Bayesian brain'. In Fotopolu, Pfaff, and Conway (eds), *From the Couch to the Lab: Psychoanalysis, Neuroscience and Cognitive Psychology in Dialogue*. Oxford: Oxford University Press.

Hopkins, J. (2018/2014). 'Kantian Neuroscience and Radical Interpretation: Ways of Meaning in the Bayesian Brain'. In G/ Oritz-Milan and J. Parero (eds), *Mind, Language, and Morality*, London: Routledge.

Hopkins, J. (2015). 'The Significance of Consilience: Psychoanalysis, Attachment, Neuroscience, and Evolution'. In S. Boag, L. Brakel, and V. Talvete (eds), *Psychoanalysis and Philosophy of Mind: Unconscious mentality in the 21st century*. London: Karnac.

Hopkins, J. (2016). 'Free Energy and Virtual Reality in Psychoanalysis and Neuroscience: A Complexity Theory of Mental Disorder'. *Frontiers in Psychology: Frontiers in Cognitive Science 7*: 922.

Hopkins, J. (2018a). 'Darwin, Freud, and Group Conflict'. In K. Figlio and S. Krueger (ed.), *Fomenting Political Violence* London: Polity Press.

Hopkins, J. (2018b). 'Freud and Mental Health'. In P. Adamson (ed.), *The Oxford Handbook of the Philosophy of Health*. Oxford: Oxford University Press.

Horton, C. L. and Malinowski, J. (2015). 'Autobiographical memory and hyperassociativity. *Frontiers in Psychopathology 6*: 874. doi:10.3389/fpsyg.2015.00874

Howe, G. (2011). *Attachment Across the Lifecourse*. New York: Palgrave Macmillan

Johnson, S., Amso, D., and Slemmer., J. (2003). 'Development of object concepts in infancy: Evidence for early learning in an eye-tracking paradigm'. *PNAS 18*(100): 10568–10573.

Johnson-Laird, P. (1983). *Mental Models*. Cambridge: Cambridge University Press.

Kant, I. (1764/2007). 'Essay on the maladies of the head'. In P. Guyer and H. Wood, *The Cambridge Edition of the Works of Immanuel Kant in Translation: Anthropology, History, and Education*, trans. H. Wilson. Cambridge: Cambridge University Press.

Kaufman, J., Csibra, G., and Johnson, M. H. (2005). 'Oscillatory activity in the infant brain reflects object maintenance'. *Proceedings of the National Academy of Science USA 102*: 15271–15274. doi:10.1073/pnas.0507 626102

Klein, M. (1952/1972). 'The Origins of Transference'. In *Envy and Gratitude and Other Works 1946–1953*. London: Hogarth Press.

Kumin, I. (1996). *Pre-object Relatedness: Early Attachment and the Psychoanalytic Situation*. New York and London: Guilford Press.

Kurth, S., Olini, N., Huber, R. et al. (2015). 'Sleep and Early Cortical Development'. *Current Sleep Medicine Reports 1*(1): 64–73. doi:10.1007/s40675-014-0002-8

Latchoumane, C., Ngo, H., Born, J., and Shin, H. (2017). 'Thalamic Spindles Promote Memory Formation during Sleep through Triple Phase-Locking of Cortical, Thalamic, and Hippocampal Rhythms'. *Neuron 95*(2): 442–435. doi:10.1016/j.neuron.2017.06.025

LeDoux, J. (2015). *Anxious: The Modern Mind in the Age of Anxiety*. London: Oneworld Publishers.

LeDoux, J. and Brown, R. (2017). 'A higher-order theory of emotional conscousness'. *PNAS 114*(10). doi:10.1073/pnas.1619316114

Leung, S., Mareschal, D., Rowsell, R., Simpson, D., Iaria, L., and Kaufman, J. (2017). 'Oscillatory Activity in the Infant Brain and the Representation of Small Numbers'. *Frontiers in Systems Neuroscience: 10*(4).

Levin, R. and Nielsen, T. (2009). 'Nightmares, bad dreams, and emotion dysregulation'. *Current Directions in Psychological Science 18*: 84–88. doi:10.1111/j.1467–8721.2009.01614.x

Lewis, P., Knoblich, G., and Poe, G. (2018). 'How memory replay in sleep boosts creative problem-solving'. *Trends in Cognitive Sciences 22*: 6. doi:10.1016/j.tics.2018.03.009

Li, W., Ma, L., Yang, G., and Gan, W.-B. (2017). 'REM sleep selectively prunes and maintains new synapses in development and learning'. *Nature Neuroscience 20*: 427–437. doi:10.1038/nn.4479

Malinowski, J. and Horton, C. (2014a). 'Emotion but not stress modulates the incorporation of waking experiences into dreams'. *Dreaming 24*: 18–31. doi:10.1037/a0036017

Malinowski, J. and Horton, C. (2014b). 'Memory sources of dreams: The incorporation of autobiographical rather than episodic experiences'. *Journal of Sleep Research 23*: 441–447. doi:10.1111/jsr.12134

Malinowski, J. and Horton, C. (2015). 'Metaphor and hyperassociativity: The imagination mechanisms behind emotion assimilation in sleep and dreaming'. *Frontiers of Psychology 6*: 1132. doi:10.3389/fpsyg.2015.01132

McCrory, E., Gerin, M., and Viding, E. (2017). 'Annual Research Review: Childhood maltreatment, latent vulnerability and the shift to preventative psychiatry—the contribution of functional brain imaging'. *Journal of Child Psychology and Psychiatry 58*(44): 338–357.

Metzinger, T. and Weise, W. (2017). *Philosophy and Predictive Processing*. Frankfurt am Main: MIND Group.

Mosquera, D., Gonzales, A., and Leeds, A. (2014). 'Early experience, structural dissociation, and emotional dysregulation in borderline personality disorder: The role of insecure and disorganized attachment'. *Bordeline Personality Disorder. Emotional Dysregulation 1*: 15. doi:10.1186/2051-6673-1-15

Panksepp, J. (1998). *Affective Neuroscience*. Oxford: Oxford University Press.

Parsons, C., Young, K., Joensson, M., Brattico, E., Hyam, H., Stein, A., Green, A., Aziz, T., and Kringlebach, M. (2014). 'Ready for action: A role for the human midbrain in responding to infant vocalizations'. *SCAN 9*: 977–984.

Piaget, J. and Inhelder, B. (1969). *The Psychology of the Child*. London: Routledge and Kegan Paul.

Pellouzio, G., Rigoli, F., and Friston, K. (2015).'Active Inference, homeostatis regulation, and adaptive behavioural control'. *Progress in Neurobiology 134*: 17–35. doi:10.1016/j.pneurobio.2015.09.001

Ramirez, S., Liu, X., MacDonald, C., Moffa, A., Zhou, J., Redondo, R., and Tonegawa, S. (2015). 'Activating positive memory engrams suppresses depression-like behaviour'. *Nature 522*: 336–339.

Ramstead, M., Badcock, P, and Friston, K. (2017). 'Answering Schroedinger's Question: A Free-Energy Formulation'. *Physics of Life Review 24*: 1–16. doi:10.1016/j.plrev.2017.09.001

Rasch, B. and Born, J. (2013). 'About sleep's role in memory'. *Physiological Reviews: American Journal of Physiology 93*: 681–766.

Reissland, N., Francis , B., Mason, J. and Lincoln, K (2011). 'Do Facial Expressions Develop before Birth?'. *PLoS ONE 6* (8): e24081. doi:10.1371/journal.pone.0024081

Reissland, N., Francis, B., and Mason, J. (2013). 'Can Healthy Fetuses Show Facial Expressions of "Pain" or "Distress"?'. *PLoS ONE 8*(6): e65530. doi:10.1371/ journal.pone. 0065530

Saarimäki, H., Gotsopoulos, A., Jääskeläinen, I., Lampinen, J., Vuilleumier, P., Hari, R., Sams, M., and Nummenmaa, L. (2016). 'Discrete Neural Signatures of Basic Emotions'. *Cerebral Cortex 26*(6): 2563–2573. doi:10.1093/cercor/bhv086

Schaffer, H. R. (1971). *The Growth of Sociability*. Harmondsworth: Penguin Books.

Segal, H. (1964). 'Symposium on Phantasy—Phantasy and Other Mental Processes'. *International Journal of Psycho-Analysis 45*: 191–194.

Segal, H. (1975/1981). 'A psycho-analytic approach to the treatment of schizophrenia'. In *The Work of Hanna Segal* (pp. 131–136). New York: Jason Aronson.

Seth, A. (2015). 'The Cybernetic Bayesian Brain'. In T. Metzinger and J. M. Windt (eds), *Open MIND*: 35(T). Frankfurt am Main: MIND Group. doi:10.15502/9783958570108

Seth, A. and Friston, K. (2016). 'Active interoceptive inference and the emotional brain'. *Philosophical Transactions of the Royal Society 371* (1708). doi:10.1098/rstb.2016.0007

Solms, M. (2013). 'The conscious Id'. *Neuropsychoanalysis 15*(1).

Solms, M. and Panksepp, J. (2012). 'The "Id" Knows More than the "Ego" Admits: Neuropsychoanalytic and Primal Consciousness Perspectives on the Interface Between Affective and Cognitive Neuroscience'. *Brain Science 2*: 147–175. doi:10.3390/ brainsci2020147

Solms, M. and Turnbull, O. (2002). *The Brain and the Inner World: An Introduction to the Neuroscience of Subjective Experience*. London: Karnac.

Solomon, J. and George, C. (1999). *Attachment Disorganization*. New York: Guilford Press.

Spillius, E., Milton, J., Garvey, P., Couve, C., and Steiner, D. (2011). *The New Dictionary of Kleinian Thought*. Hove: Routlege.

Stenberg, C. R., Campos, J., and Emde, R. N. (1983). 'The facial expression of anger in seven-months-old infants'. *Child Development 54*: 178–184.

Stenner, K. (2010). *The Authoritarian Dynamic*. Cambridge: Cambridge University Press

Swanson, L. R. (2016). 'The predictive processing paradigm has roots in Kant'. *Frontiers in Systems Neuroscience 10*: 79. doi:10.3389/fnsys.2016.00079

Tappenberg, T. (2010). *Foundations of Computational Neuroscience*. Oxford: Oxford University Press.

Tonegawa, S., Liu, X., Ramirez, S., and Redondo, R. (2015a). 'Memory Engram Cells Have Come of Age'. *Neuron 85*: 918–931.

Tonegawa, S., Pignatelli, M., Roy, D., and Ryan, T. (2015b). 'Memory Engram Storage and Retrieval'. *Current Opinion in Neurobiology 35*: 101–109.

Tononi, G. and Cirelli, C. (2014). 'Sleep and the price of plasticity: From synaptic and cellular homeostasis to memory consolidation and integration'. *Neuron 81*: 12–34. doi: 10.1016/j.neuron.2013.12.025

Travaglia, A., Bisazi, R., Sweet, E., Blitzer, R., and Alberini, C. (2016). 'Infantile amnesia reflects a developmental critical period for hippocampal learning'. *Nature Neuroscience 19*: 1225–1233.

Vogel, G., McAbee, R. J., Barker., K., and Thurmond, R. (1977). 'Endogenous Depression Improvement and REM Pressure'. *Archives of General Psychiatry 34*: 96–97.

Volkow, N., Wise, R., and Baier, R. (2017). 'The dopamine motive system: Implications for drug and food addiction'. *Nature Reviews Neuroscience 18*(12): 741–752. doi:10.1038/nrn.2017.130

Walker, M. and van der Helm, E. (2009). 'Overnight Therapy? The Role of Sleep in Emotional Brain Processing'. *Psychological Bulletin 135*(5): 731–748.

Watt, D. and Panksepp, J. (2009). 'Depression: An evolutionarily conserved mechanism?'. *Neuropsychoanalysis 2*(1): 93.

Wilcox, T. (1999). 'Object individuation: Infants' use of shape, size, pattern, and color'. *Cognition 72*: 125–166.

Wilcox, T. and Biondi, M. (2015). 'Object processing in the infant: Lessons from neuroscience'. *Trends in Cognitive Sciences 19* (7): 406–413.

Notes:

(1.) Thus as briefly discussed later below the psychoanalytic mechanisms of identification and projection seem clearly to play a role in the ingroup cooperation for outgroup conflict that Darwin and contemporary theorists of culture gene co-evolution regard as formative for our species, and both the Oedipus complex and the death drive also seem clearly related to parent–offspring, sexual, and sibling conflict. These are discussed together in Hopkins (2018a); and Hopkins (2015) discusses the relations of psychoanalysis, evolution, attachment, and neuroscience more generally. The integration of Freudian and contemporary free energy neuroscience in the complexity theory of mental disorder sketched later on is treated at greater length in Hopkins (2016).

(2.) The programme has a (technical) website at https://en.wikipedia.org/wiki/ Free_energy_principle.

(3.) On generative models in contemporary neuroscience, see Ch. 10 of Dayan and Abbot (2001); for an exegesis of Friston's use, see Ch. 10.3ff of Tappenberg (2010).

(4.) This account was ignored by the founders of twentieth-century analytical philosophy. Succumbing to what Wittgenstein later described as 'the picture that forces itself on us at every turn', Frege, Russell, and the early Wittgenstein turned their backs on Helmholtz's conception of perceptual experience as synthesized by the brain. Instead they adopted a Cartesian conception in which such experiences were the private 'sense-data' upon which all knowledge, including the existence of bodies and brains, had somehow to be based. See Hopkins (2014/2018c) for an account of the development of Helmholtz's ideas concerning modelling through Frege and the early Wittgenstein, which ideas prompted Craik's (1943) account of mental models, later taken up in cognitive science, e.g. in Johnson-Laird (1983).

(5.) This holistic interpretive integration of the data of free association is discussed in some detail in Hopkins (1999). Although other forms of investigation of the emotional functions of dreaming increasingly acknowledge the evidential importance of the dreamer's associations and memories (see, e.g. Malinowski and Horton 2015), no mode of study apart from psychoanalysis has taken them so thoroughly into account. If as argued later dreaming functions as part of the emotional synthesis of long-term memory then progress in understanding both dreaming and memory will partly depend upon other disciplines following this lead.

(6.) In arriving at this account Freud started to provide detailed empirical support for a tradition of thought about dreaming and disorder that went back to Plato, and was summarized in Kant's (1764/2007: 71) claim that 'the deranged person' was 'a dreamer in waking'.

(7.) While still a medical student Freud was invited by the celebrated physiologist Ernest Bruke to conduct neurological research in his laboratory. Prior to practising as a psychiatrist he published over a hundred papers, as well as monographs on disorders of movement and childhood cerebral palsy that established him as an expert in these fields.

(8.) Helmholtz's notion of thermodynamic free energy measures a physical quantity, whereas the notion of variational free energy (FE) provides a measure of Bayesian model evidence that is computable from parameters of a model. Similarities and differences between the notions are discussed in some detail at https://en.wikipedia.org/wiki/ Free_energy_principle. The two are related mathematically, since the equations describing variational FE derive from Helmholtz's describing thermodynamic free energy; and for this reason they share the terminology of statistical physics. Also the complexity term of variational free energy shares the same fixed point as Helmholtz free energy for a system that is thermodynamically closed but not isolated.

(9.) 'Optimally': that is, by the criteria imposed by evolution, which would determine the prior hypotheses and major modes of processing of the model, as well as by the environment, including experiential learning over the lifespan. Hence only in some areas (e.g. perception) does the brain appear as operating with Bayesian rationality. In others (e.g. the assignment of properties to members of idealized ingroups as opposed to derogated outgroups) it must be understood as operating in accord with priors imposed by evolution (in this case those that sustain the process of ingroup cooperation for outgroup conflict, 'the competition of tribe with tribe', that Darwin thought 'sufficed to raise man to his present high position in the organic scale' (1871: 157)). In enabling us to understand such oppositions between the imperatives of rationality and evolution the FE programme encourages enquiry that parallels Quine's conception of the naturalization of epistemology.

(10.) Since the turn of the century Friston and others have directed this approach to the rest of organic life (Friston 2013; Ramstead et al. 2017) and applied it to the brain in rapidly increasing scope and detail. This has established the minimization of FE as a paradigm for neuroscience, described variously as predictive coding, predictive processing, the predictive brain, and the predictive mind. And since the probabilistic regulatory generative models envisaged in this approach would encompass a host of others—the mental models of cognitive science, the internal working models of attachment theory, the internal objects of psychoanalysis, etc.—the programme naturally extends through psychology as well.

Although Friston's own presentations are often forbiddingly technical, the main ideas have been given clear philosophical expositions in Hohwy (2013) and Clark (2016), as well as a recent collection of essays edited by Metzinger and Weise (2017). Seth (2015) is a good introductory starting point, and emphasizes the role of interoception in the generation of emotion, which is discussed later in this chapter.

(11.) This continuity contradicts claims that human thought cannot be understood as having arisen via natural selection. And although specifics of Friston's account of the brain's model are controversial, the existence of some such model (map, etc.) is widely accepted in neuroscience. The brain viewers at http://gallantlab.org/index.php/brain-viewer/ display a striking range of information about the modelling of the external world in the cortex (see also the video introduction at https://www.youtube.com/watch?v=u9nMfaWqkVE).

(12.) Friston's collaborator Alan Hobson, for decades a zealous neuroscientific critic of Freud, has now published two books (Hobson 2014, 2015) applying the virtual reality/generative model conception to link dreaming with mental disorder, although via neurotransmission rather than Freudian phantasy. While continuing to criticize Freud, Hobson now stresses that his own work 'takes up the *Project for a Scientific Psychology* exactly where Freud left it in 1895' (Hobson 2015: 5).

(13.) These systems are described differently by different investigators. For example, LeDoux (2015) describes them as 'survival circuits', and Volkow et al. (2017) describe the SEEKING system as the Dopamine Motive System. The most detailed recent delineation is Panksepp's (1998), which has been updated for therapists and general readers in Bivens and Panksepp (2011).

(14.) One important difference is that Freud did not anticipate the current claim that the brainstem systems generate consciousness as well as emotion, both of which are elaborated in the cortex. For this see Solms and Panksepp (2012). Again many authors, such as Barrett (2017) and LeDouz and Brown (2017) argue that the contextual cortical elaboration of emotion is inconsistent with the notion of basic emotions. But as Panksepp argues, subcortical generation of the basics is consistent with any degree of cortical elaboration of the final states, including that emphasized by Barrett; and as noted by Saarimäki et al. (2016) the subcortical generators have 'fingerprints' in their complex cortical derivatives.

(15.) This is the kind of dialectic that Bion (1962: 30) attempts to describe phenomenologically as follows:

> The infant suffers pangs of hunger and feels it's dying; racked by guilt and anxiety and impelled by greed, it messes itself and cries. The mother picks it up, feeds it and comforts it and eventually the infant sleeps. In forming the model to represent the feelings of the infant, we have the following version: the infant, filled with painful lumps of faeces, guilt, fears of impending death, chunks of greed, meanness and urine, evacuated these bad objects into the breast that is not there. As it does so, the good object turns the no-breast (mouth) into a breast, the faeces and urine into milk, the fears of impending death and anxiety into validity and confidence, the greed and meanness into feelings of love and generosity; and the infant sucks in its bad property, now translated into goodness, back again.

(16.) Some of main developments occur in and between the 'what' and 'where' object-processing pathways (Wilcox and Bondi 2015), with an early representation of an identified object as lasting while unobserved *perhaps* signalled at six months by gamma wave activation in the right temporal cortex (Kaufman et al. 2005; see also Leung et al. 2017).

This is consistent with an experimental tradition that at first focused on the mother as object of perception, but later on other (emotionally less significant) objects. Thus Bower (1977) reported an experiment with mirrors (discussed in Hopkins 1987) that cohered with Klein's dating of the depressive position.

> If one presents the infant with multiple images of its mother—say three 'mothers'—the infant of less than five months is not disturbed at all but will in fact interact with all three 'mothers' in turn. If the setup provides one mother and two strangers, the infant will preferentially interact with its mother and still show no

signs of disturbance. However, past the age of five months (after the coordination of place and movement) the sight of three 'mothers' becomes very disturbing to the infant. At this same age a setup of one mother and two strangers has no effect. I would contend that this in fact shows that the young infant (less than five months old) thinks it has a multiplicity of mothers, whereas the older infant knows it has only one

(Bower 1977: 217).

Although this experiment has not been repeated in contemporary conditions, it was consistent with experiments on occlusion (e.g. Baillargeon et al. 1985; Wilcox 1999) and eye-tracking (Johnson et al. 2003) done afterwards. Later work (Baillargeon et al. 2012) has clarified how the infant initially uses parameters of shape and then others as the ability to individuate and track objects is refined.

(17.) See also Klein (1952/1972: 54). For comparison of Piaget and Klein on infantile experience and the object concept see Hopkins 1987.

(18.) Thus as reported in (Campos et al. 1983), a three-month-old infant made angry by someone using her hand to impede movement will express anger *at the impeding hand*. By contrast a seven-month-old infant made angry in the same way expresses anger *not at the hand, but at the face*. Apparently at three months the infant has not, whereas by seven months it has, integrated episodic experiences with the mother into an image of an anatomically whole and enduring person. Also by seven months, as reported in (Stenberg et al. 1983) this infant's anger is regulated by experience of the persons she knows. If the infant is annoyed twice by mother, or again twice by a stranger, the infant is angry both times. But an infant annoyed *first* by a stranger and *then* by mother is especially angry, indicating that the infant's expectation that mother will comfort her after intrusions by a stranger has been betrayed.

(19.) The work reported in Ali et al. (2015) suggests that infants' sense of bodily touch becomes anchored to objects in the surrounding space at four to six months. And recently Seth and Friston (2017) have suggested that failure in integrating interoceptive and exteroceptive sensory information—in particular, failure properly to integrate 'interoceptive signals associated with nurturance (e.g. breastfeeding) during affiliative interactions with (m)others' may render autistic infants incapable of learning 'that the nursing and prosocial mother were the same hidden cause or external object' (Seth and Friston 2017: 7, and Fig. 4). If this is correct it would place autism among the disorders that result from failure to negotiate this depressive-position integration.

(20.) For a fuller discussion of these developments, which also include the regulation of parent–offspring conflict and sibling rivalry as part of the Oedipus complex, see Hopkins (2018b).

(21.) For further discussion and citations see the section on 'Waking Disorder and Complexity Reduction in Sleep' in Hopkins (2016) as well as the work on dreaming and depression cited later in this section.

(22.) For a fuller discussion of this dream which bears out claims made in passing here see Hopkins (2015), and for further methodological discussion see Hopkins (1999).

(23.) See also Walker and van der Helm (2009) and Cartwright (2010).

(24.) On symbolism see Petocz, this volume. For the symbolism in this dream see the discussions in Hopkins (2015), and Hopkins (2002) on psychoanalysis and conceptual metaphor, which is also discussed in the work by Malinowski and Horton cited here.

(25.) See the discussion in Cartwright (2010: ch. 4), following up the work Vogel et al. (1977); and the continuation of Cartwright's work in the neuropsychoanalytic research project on dreaming and depression described in Fischmann et al. (2013).

(26.) Thus for example the main arguments and claims of a recent series of non-psychoanalytic papers by Malinowski and Horton (e.g. 2014a, 2014b, 2015; see also Horton and Malinowski 2015) are consilient both with Freud and with the argument of the present paper. They discuss role of memory, the dreamer's associations, and symbolism, and see dreams as performing 'emotion assimilation' via symbolic embodied cognition that ameliorates the 'emotional intensity' of memories transformed in them.

Jim Hopkins

Department of Philosophy, King's College London

Oxford Handbooks Online

Abstract and Keywords

Acknowledging the layers of the mind below the level of overt consciousness can lead to very divergent accounts of religious belief. One response—taken by Freud himself—argues that religious belief should be abandoned as unavoidably contaminated by unconscious motivations (e.g. an infantile longing for security) that distort our rational judgement. By contrast, Jung maintains that religious thinking is shaped by unconscious structures (the 'archetypes'), which can play a vital role in the development of an integrated human personality. This chapter examines these contrasting psychoanalytic interpretations of religion, and then explores more recent accounts of the workings of the human psyche and how they affect the status of religious belief. A concluding section discusses some general implications of all this for the epistemology of religious belief and the way in which philosophy of religion should be conducted.

Keywords: Freud, Jung, unconscious, illusion, archetypes, Graham Ward, Iain McGilchrist, brain hemispheres, receptivity, detachment

Introduction

A large part of the theory and practice of psychoanalysis, from Sigmund Freud onwards, has been concerned with the treatment of disturbed and troubled patients seeking help with their symptoms, but the philosophical importance of the Freudian revolution is very far from confined to the domain of the pathological or the neurotic. In effect what Freud challenged was a simplistic but widely held model of the mind as a kind of transparent goldfish bowl inside which straightforwardly identifiable beliefs and desires float around, ready to inform our actions and choices. In the wake of Freud, it has become much harder to be confident about the image of ourselves as self-sufficient and autonomous rational agents whose decisions are based solely on what is straightforwardly accessible to the reflective mind. This does not mean that the Freudian revolution in our conception of ourselves undermines the very possibility of rational thought—if that were the case, psychoanalytic thought would be self-refuting, since it would undermine the possibility of its coherent articulation. What *is* entailed is that we should give up the naïve conception of our mental powers and capacities as transparent tools of reason, and start working towards a more nuanced conception, according to which uncovering the truth about ourselves and our relation to the world must be approached in a spirit of humility and receptivity that acknowledges the intensely complex and problematic nature of the instrument which we must use to undertake that task—the human mind.

These general lessons of the Freudian revolution evidently have application to the domain of religious belief, along with many other areas of human thought. But acknowledging the layers of the mind that operate below the level of overt consciousness can lead to very divergent accounts of the status and validity of religious beliefs and attitudes. One way to go—the route that Freud himself took—is to argue that religious belief should be abandoned, insofar as it is unavoidably contaminated by unconscious drives and motivations (an infantile longing for security, for example) that distort our rational judgement. A quite opposite approach, exemplified by that of Freud's onetime disciple Carl Jung, is to maintain that religious thinking is typically shaped by unconscious forms and structures (what Jung called the 'archetypes') which, so far from being generators of neurosis, can play a vital role in the development of a healthy and integrated human personality. We shall look in more detail at these two influential but strongly contrasting psychoanalytic interpretations of religion in the next two sections, before going on, in the fourth section, to explore more recent accounts of the workings of the human psyche and how they may affect the status of religious belief. The fifth and final section will aim to tease out some general conclusions about the relationship between psychoanalysis and religion, and the implications of this for the epistemic status of religious belief, and the way in which the philosophy of religion should be conducted.

Freud's Critique of Religion

Freud's attack on religion begins by drawing attention to our human helplessness before the 'majestic, cruel and inexorable powers of nature' (Freud 1927: 195). These powers include both external forces (earthquakes, floods, hurricanes) and the equally threatening internal forces (lust, anger, brutality) arising from our own nature. Freud sees religion as an attempt to mitigate our defencelessness by endeavouring to 'adjure, appease, bribe' or otherwise influence a celestial father figure, who will protect us from suffering, and impose justice on a seemingly chaotic and terrifying universe (1927: 196).

The vulnerability of the human condition, and the fact that since time immemorial humans beings in extremis have resorted to a variety of supposed divine powers and forces to rescue them when all else fails, are familiar enough themes which have been commented on by many writers, including Jean-Paul Sartre:

> When the paths traced out become too difficult, or when we see no path, we can no longer live in so urgent and difficult a world. All the ways are barred. However, we must act. So we try to change the world, that is, to live as if the connection between things and their potentialities were not ruled by deterministic procedures, but by magic
>
> (Sartre 1975: 58–9).[1]

The same general line had been taken much earlier by another stern critic of religion, David Hume. What prompts us to suppose there is a God, according to Hume, are 'the ordinary affections of human life' such as the 'dread of future misery' and the 'terror of death' (Hume 1757: section 2).

But to explain the religious impulse simply in terms of human vulnerability and helplessness leaves something out: no doubt people earnestly desire to be rescued when in trouble, but that in itself does not seem to account for the strength and pervasiveness of the human belief in the divine. When in grave distress we might *like* to deceive ourselves into thinking the world is determined not by natural forces but by magic (as Sartre phrases it), but this does not in itself explain how widespread and successful such self-deception (if it is indeed that) has become. Here Freud contributes something crucially important by addressing himself to the psychological question of *how* the belief becomes so powerfully entrenched in the minds of so many religious adherents. The human psyche, he argues in *The Future of an Illusion*, is *already predisposed*, as a result of the traces left by our forgotten experience as infants, to conjure up the image of a powerful protector to rescue us from our helplessness.[2] For when we encounter threats and dangers:

… this situation is nothing new. It has an infantile prototype of which this is in fact only the continuation. For once before one has found oneself in a similar state of helplessness: as a small child, in relation to one's parent. One had reason to fear them, and especially one's father; and yet one was sure of his protection against the dangers one knew. Thus it was natural to assimilate the two situations. Here, too, wishing plays its part, as it does in dream-life …

(Freud 1927: 196).

So just as with the strange deliverances of dreams, what is planted in our consciousness has a resonance, a power that takes hold of us quite independently of the normal criteria of reasonable evidence and rational judgement. The mind is in the grip of an *illusion*, but this is not just a mistake, or a piece of deliberate self-deception. Rather, layers of mentation working beneath the level of explicit awareness or rational reflection have been activated by our helplessness in the face of the perils we face as adults, and we revert, without being consciously aware of what is going on, to the infantile state of fear and dependency which is ineradicably linked to the yearning for security and the hope of parental protection. Only as a result of delving into the deeper workings of the mind right back from early childhood does the full explanation of the process come to light. This is the background that enables Freud to declare with such confidence 'the derivation of religious needs from the infant's helplessness and the longing for the father aroused by it seems to me incontrovertible' (Freud 1929: 260). So we arrive at the famous Freudian diagnosis: religion is an illusion born of helplessness and fear.

Freud makes it clear, however, that he does not intend his psychoanalytic diagnosis of our longing for protection to be a logical demonstration of the falsity of the religious world view. That would be to commit the 'genetic fallacy' (the logical fallacy of confusing the causal origins of a belief with its justification or lack thereof). Illusions, as Freud concedes, are not *necessarily* erroneous: 'A middle-class girl may have an illusion that a prince will come and marry her … and a few such cases have actually occurred' (Freud 1927: 213). But Freud argues that it is characteristic of illusions in his sense that they are held on to without regard for rational justification; further, they characteristically stem from (indeed are generated by) the wishes or needs of the believer. So it is a short step from this to the conclusion that Freud is aiming at: religion is an infantile piece of wishful thinking that we need to grow out of.

Yet on further reflection the implications of Freud's critique are by no means as damaging as might at first appear. The believer might well concede to Freud that our infantile helplessness leaves a lasting stamp on the psyche, but go on to insist that this can scarcely be the whole story. For beyond any mere desire for protection (Freud's 'longing' for the father figure), it seems hard to deny that the religious impulse is in large part connected with the powerful yearning human beings have for meaning and purpose in their lives. Now it could be proposed, as is done by many secularists, that meaning and purpose must be found in the chosen activities and pursuits—intellectual, artistic, social, familial, and so on—which are the components of a worthwhile human life. But, without

denying the value and meaningfulness of such activities and pursuits, it may be argued that they cannot in themselves bear all the weight of satisfying our human hunger for meaningfulness. One way of putting this is to say that to be human is to have a characteristic restless, a sense of incompleteness, such that even were all our specific needs and goals to be satisfied (for food, for shelter, for company, for recreation, for satisfying relationships, for creative activities, and so on), there would always remain a longing for something more—something that will provide an 'ultimate grounding' for our lives, or give us a sense of 'ontological rootedness' (May 2011: 7).

God, for the religious believer, is the ultimate source of being and value towards which we yearn, and which alone can satisfy the existential longing which is part of the nature of dependent and contingent beings such as us.[3] Pointing this out does not of course vindicate belief in God, nor does it of itself refute deflationary Freudian-style explanations of it, but it may at least open up the possibility that religious belief connects with something in our human nature of deeper significance than a mere neurotic or infantile impulse. Certainly there are many places in Scripture where the strange open-ended longing of the human spirit is underlined ('Like as the hart desireth the water-brooks, so longeth my soul after Thee, O God' (Psalm 42 [41]: 1)); and the theme is reiterated in seminal Christian writers such as St Augustine ('You have made us for Yourself, and our heart is restless until it finds repose in You', and Dante ('In his will is our peace')).[4] The thought in such passages is not merely that religious devotion provides peacefulness of mind, in the sense of securing some kind of tranquillizing or calming effect; rather, the idea is that God is the source of genuine value, and that orienting ourselves towards that source bestows ultimate meaning on our human existence and enables us to find true fulfilment even in the face of danger and turmoil. Augustine and Dante acknowledge our vulnerability, but manage to construe it as a corollary of our creatureliness, so they can end up celebrating it as a cause for joyful affirmation of our creator. Freud by contrast sees our vulnerability as a condition which scares us so much that we desperately fantasize that we have found a way of assuaging it—even though in fact the power we appeal to has no reality outside the human psyche. But as to which of the two accounts reflects the way things actually are, this remains to be determined; so however disconcerting Freud's analysis may initially be for the religious believer, it seems clear that it cannot finally settle the matter.

Carl Jung and the Importance of Symbolic Thought

While Freud's view of religious belief places it under the same general heading as the neurotic, the infantile, and the disturbed, or at any rate as falling short of the standards of balanced judgement to which we aspire as reasonable adults, Carl Jung, though concurring with Freud that such belief has roots buried deep in the human psyche, took a very different view of the resources of the unconscious mind, regarding them in a potentially much more benign light.[5] For Jung, Freud's dismissal of the religious impulse as infantile fails to recognize the imaginative and symbolic role of religious modes of thought and expression, and their possible role in the healthy development of the human personality. Crucial here is the idea of 'individuation' as Jung terms it, the 'process by which a person becomes a psychological "in-dividual", that is, a separate, indivisible unity or whole' (Jung 1939: 275). Jung sees human psychological development in terms of a struggle to achieve internal balance and psychic integration, where integrating the conscious and unconscious elements of the self is a precondition for psychic health or wholeness;[6] and religious imagery and symbolism, according to Jung, perform a vital function here (Jung 1952). The process of individuation requires modes of thought and expression that operate not just on the surface level of explicit assertion, but which carry deep imaginative resonances that are vital for our psychological balance and harmony. To give but one example of this, the figure of Christ functions for Jung as an 'archetype of the self', a deeply resonant image of the perfectly unified and integrated human being (Jung 1938; Jung 1951: 183). From this perspective, as Michael Palmer aptly puts it in his account of the Jungian position:

> Religion, far from being neurotic, is revealed as a constant and evolving process in the development of the psychic personality … Religious symbols … open up a psychic level … that is primordial and … of supreme value for the present and future development of the human psyche
>
> (Palmer 1997: 110–11).

Jung's ideas have encountered considerable philosophical opposition (as indeed have those of Freud). Many contemporary analytic philosophers are supporters of what Brian Leiter has called the 'naturalistic revolution in philosophy', according to which philosophy should 'adopt and emulate the methods of the successful sciences' (Leiter 2004: 2–3). Such philosophers often tend to be sceptical about the very idea of the unconscious mind, and *a fortiori* the Jungian idea of the archetypes, on the grounds that the theories that invoke such ideas lack the kind of hard scientific warrant demanded by today's dominant naturalistic paradigm. In response to this kind of critique, defenders of the Jungian approach have two possible lines of defence. One is to argue that there is in fact hard empirical evidence, for example from cognitive science and developmental psychology, that can be used to support the Jungian hypothesis of the role of symbols and archetypes

in psychic integration (Knox 2003). Another response would be to take issue with the doctrine that the methods of science are the only valid way of uncovering the truth. Thus Thomas Nagel has argued, in the case of Freud, that irrespective of what we think about the authority of analysts or the clinical evidence for their theories, there is an 'evident usefulness of a rudimentary Freudian outlook in understanding ourselves and other people, particularly in erotic life, family dramas, and what Freud called the psychopathology of everyday life' (Freud 1901; Nagel 1994). And similarly one could argue that Jung's ideas, like those of Freud, are best assessed not as contributions to science, but in a more 'hermeneutic' way—that is, as ways of enriching our understanding of the human predicament, and the deeper significance of our thoughts and feelings and beliefs.[7], [8]

To suggest that Jung's ideas may enrich our understanding without qualifying as contributions to science is not at all to dismiss or downgrade the value and importance of scientific inquiry or scientific methods. One can be a genuine and wholehearted admirer of the achievements of *science* while at the same time resisting the false allure of *scientism*—the dogma that scientific methods give us everything we need to understand all aspects of reality. To be sure, we live in, and are an integral part of, the physical world constituted by the particles and forces studied by science—that is undeniable. But when it comes to understanding aspects of human life such as religious experience (and the same goes for poetic or artistic or moral experience, or even our ordinary human interactions with each other) we patently need other categories than those of the physical sciences; for even the fullest and most detailed print-out of the relevant particle collisions and biochemical processes will tell us nothing about the *human significance* of these processes and events.[9]

In addition to the physical sciences there are of course the social sciences (including for example economics, sociology, and psychology), and there are continuing debates as to how far such disciplines meet the standards of the 'hard' physical sciences (in matters such as experimental repeatability, verifiable prediction, mathematical modelling, and so on). There is probably no simple answer to this question, since the term 'social science' covers a large array of divergent disciplines and inquiries, whose methods manifest varying degrees of precision and rigour. A particular issue as regards psychology is the inevitable reliance on reports by individual human subjects of their thoughts, feelings, sensations, beliefs, and desires, thus making reference to what is, according to some philosophers, an irreducible domain of qualitative subjective experience that resists subsumption or explanation in objective scientific terms (Nagel 1974). But however that may be, psychoanalytic approaches to psychology present special additional problems, insofar as the 'data' being studied come to light in the context of special relationship with the analyst—an issue which led even Freud, despite his attraction to the scientific model, to remark on how far psychoanalysis diverges from normal scientific procedures. We shall return to this issue in 'Philosophizing about Religion and the Layers of the Human Psyche'.

At all events, when we come to religious feelings, beliefs, attitudes, and experiences, it is apparent that these have a characteristic depth and complexity that can seldom if ever be conveyed in a set of straightforward factual propositions laid out for our assessment and awaiting verification. For religious thoughts and ideas operate within a rich and complex web of *associations*, carrying manifold metaphorical and symbolic echoes which may often have powerful effects on us in ways that are not fully transparent to consciousness. This, as Jung sees it, is the key to the peculiar resonance and power of the images and icons that inform the thoughts and ideas of religious believers, and what explains their role in the search for integration and healing within the troubled human psyche. None of this of course means that we should uncritically accept all or any of Jung's ideas about the role of religious concepts; but at least it reminds us of the context in which his theories are meant to operate, and within which they need to be evaluated.

Whatever conclusions one finally reaches about the Jungian theory of archetypes, important questions remain about Jung's general approach to religion. The foremost among these is the objection that the Jungian approach leads to a kind of psychologizing or subjectivizing of religion, where the question of the truth or validity of any given religious outlook (Christian theism, for instance) boils down to no more than the question of whether certain archetypal images (such as that of God the Father, or Christ the Son, for example) have a transformative power within the human psyche (Palmer 1997: 187, 196). Jung's own response to this type of criticism was that his role as a psychologist was not to make pronouncements about the existence or non-existence of transcendent realities, but simply to describe the role of certain fundamental and universal images and symbols in human development:

> We know that God-images play a great role in psychology, but we cannot prove the [actual] existence of God. As a responsible scientist, I am not going to preach my personal and subjective convictions which I cannot prove … To me, personally speaking, the question whether God exists at all or not is futile. I am sufficiently convinced of the effects man has always attributed to a divine being. If I should express a belief beyond that … it would show that I am not basing my opinion on facts … I am well satisfied with the fact that I know experiences which I cannot avoid calling numinous or divine

> (Jung 1956).

This makes quite clear the restricted scope of Jung's position: it insists that religious concepts and images play a crucial role in the development of the human personality and its search for integration, but leaves completely open the question of whether there is some objective reality—something 'external' or independent of the subjective structure of the human psyche—to which those concepts and images refer.

To sum up our necessarily compressed and selective account of the contrasting attitudes of Freud and Jung to religion—the former's highly negative, the latter's much more positive—what emerges is that in neither case do their suggested findings about the workings of the human psyche in themselves either establish or refute the truth of the

religious outlook. The partly hidden motivation for religious belief may, if Freud is right, be an infantile one; but as we have seen, that in itself does not logically entail the falsity of such a belief. And the symbols and images drawn from the unconscious mind may, if Jung is right, exert a powerful psychological influence on the human quest for integration; but that, as just noted, still leaves open the real existence or otherwise of the God that is the object of religious belief.

In the wake of these two seminal thinkers, however, one thing at any rate should be clear: that any philosophical attempt to address the fundamental questions of religious belief will find it hard to carry conviction unless it takes some account of the complexity that lies beneath the seemingly transparent surface of propositional assent to religious claims and doctrines. No account of religious belief and experience is going to look plausible unless it acknowledges the complexity of the human mind—the strata of hidden longings and needs and the manifold symbolic forms and images resonating deep within the human psyche. To some of the more recent attempts to address that complexity we shall now turn.

The Complexity of Belief

The contemporary debate over the validity of religious belief tends to play out in a curiously abstract and rationalistic way. The implicit assumption is that the participants are detached evaluators, judiciously examining the 'God hypothesis' (as Richard Dawkins calls it (Dawkins 2006: ch. 2)), scrutinizing the supposed evidence, and weighing up the arguments for and against. Particularly among anglophone philosophers in the analytic tradition, where psychoanalytic ideas have tended to be resisted or ignored by many practitioners, the claims of religion are implicitly construed as rather like scientific claims, suitable subjects for purely intellectual disputation, the province, as it were, of scholarly discussion in the seminar room. But a number of recent writers have started to challenge this austere and bloodless picture of religious belief.

Philosophers have argued endlessly about the epistemic status of religious and other kinds of belief, and what entitles some beliefs to the accolade 'knowledge', but comparatively few have paid attention to 'what lies beneath'—to the 'archaeology of belief', as the British theologian Graham Ward has called it. Ward argues that believing or disbelieving something involves far more complex processes than the scrutiny and evaluation of factual evidence. There are much 'deeper layers of embodied engagement and reaction', where we are touched 'imaginatively, affectively and existentially' (Ward 2014: 7, 10, 31). Drawing on empirical research into the behavioural and neurological underpinnings of belief, and its evolutionary and prehistoric roots, Ward delves into the domain of what the Berkeley psychologist John Kihlstrom has termed the 'cognitive unconscious' (Kihlstrom 1987; see also Kahneman 2011). A rich array of non-conscious mental activity, including learned responses that have become automatic, subliminal perceptions that impact on our conscious judgements, and implicit but not consciously recalled memories—all these profoundly affect how we perceive and interpret the world (Ward 2014: 11, 68). And as we saw in the case of Freud (whose general influence is clearly discernible here), the implications of the resulting conception of human belief and understanding extend far more widely than the domain of the pathological. Not just in neurotic desires and perceptions, but whenever we believe anything at all, there is, as Ward puts it, a 'mode of liminal processing, related to embodiment and affectivity, which "thinks" more quickly and reacts more instinctively than our conscious rational deliberation' (Ward 2014: 12).

A further dimension of complexity in our beliefs is explored in the neuroscientist Iain McGilchrist's groundbreaking work *The Master and His Emissary*, according to which there are two different modes of relating to the world, broadly correlated with the activities of the left and right hemispheres of the brain respectively, one mode being detached, fragmented, abstract, and analytical, the other being more direct, holistic, intuitive, and empathetic. McGilchrist speaks of:

> … two ways of being in the world, both of which are essential. One is to allow things to be *present* to us in all their embodied particularity, with all their changeability and impermanence and their interconnectedness, as part of a whole which is forever in flux … The other is to step outside the flow of experience and 'experience' our experience in a special way: to *re-present* the world in a form that is … is abstracted, compartmentalised, fragmented, static … From this world we feel detached, but in relation to it we are powerful (2009: 93).

Acknowledging McGilchrist's influence, Graham Ward urges us to question the 'left-brain hegemony' that has increasingly dominated our culture since the Enlightenment, and to reconfigure our understanding of belief. Doing justice to the full range of our embodied human engagement with the world could allow for a 'rebalancing' of left-brain and right-brain modes of awareness. This, he argues, might enable us to overcome the sterile opposition between scientific and religious modes of thinking, and to understand what lies 'at the very core of poetic and religious faith' (Ward 2014: 110).

Before proceeding, it needs to be noted that many critics have questioned the distinction that is invoked here between 'left-brain' and 'right brain' activity, objecting that the available scientific research on the neurophysiology and functioning of the brain does not support a strong dichotomy between the functions performed by the two hemispheres (Nielsen et al. 2013). There may be evidence to suggest that in most subjects one can distinguish between 'logical–conceptual' and more 'intuitive' mental activity, each broadly correlating with neural activity in the relevant halves of the brain; but the critics point out that in normal subjects both halves play some role in both, and in any case there is constant interaction between the two hemispheres. All this, however, is readily conceded by McGilchrist, who fully acknowledges the massive degree of interconnectivity in the wiring of the brain, while nevertheless insisting that the two hemispheres have been shown to function in ways that are to some degree independent, and that this can tell us something important about the different ways in which we experience the world (McGilchrist 2018).[10]

However that may be, the position taken by McGilchrist, Ward, and others[11] about the need to challenge what they term 'left-brain hegemony' does not seem ultimately to hinge on the precise details as to how the brain is configured. For the crucial point at issue is not a neurological one, but what might be called a psycho-ethical or spiritual one: that our ultimate flourishing as human beings depends on our being able to integrate our detached and analytic modes of relating to the world with our more direct and intuitive modes of awareness. This is not to say, however, that the scientific study of the brain has no relevance to the psychological-cum-moral task of striving for an integrated vision of the world. For the wiring of the brain, shaped by the long history of its evolution, is an integral part of our nature as biological creatures, and our human ways of perceiving and understanding the world must inevitably be conditioned and mediated by that history. The point was in fact explicitly anticipated by Jung in a paper written early in his career:

> Just as the human body represents a whole museum of organs, with a long
> evolutionary history behind them, so we should expect the mind to be organized in
> a similar way … We receive along with our body a highly differentiated brain
> which brings with it its entire history, and when it becomes creative it creates out
> of this history—out of the history of mankind … that age-old natural history which
> has been transmitted in living form since the remotest times, namely the history of
> the brain structure

(Jung 1918: 12).[12]

In short, whatever scientific consensus is eventually reached with regard to the precise workings of the brain and the functioning of its parts, the resulting picture seems likely only to reinforce the idea that our grasp of reality depends at the physiological level on an intricate nexus of mechanisms and processing systems evolved over many millennia and working beneath the threshold of conscious awareness and control. And alongside this neurological complexity there also has to be taken into account the complex array of socially and culturally inherited associations and resonances that condition our cognitive and emotional responses to the world, again working largely below the level of our explicit conscious awareness. So the more we learn about all this, the more pressure is put on the idea of the detached autonomous agent, somehow operating above the fray of evolution and history, and forming beliefs based solely on dispassionate scrutiny of the evidence like some pure disembodied intelligence.

So what are the implications of the growing interest in the 'archaeology of belief' for our understanding of religion and its place in the modern scientific age? Ward, as already noted, maintains that a greater understanding of what lies beneath the surface of conscious belief-formation will help us to overcome what he sees as the 'sterile' opposition between scientific and religious thought and to give up the idea that primitive mythological ways of thinking about the world will progressively be replaced by modern scientific methods. For placing the belief-forming faculties of our species within the context of their biological and social development over many millennia reveals the ineradicable role of the mythic and the symbolic in *all* human cognition, and thus radically undermines the idea of the inevitable triumph of a science-based, demythologized and secularized belief system. Just as Jung had argued that mythical and symbolic forms powerfully and inescapably impinge on our human beliefs and attitudes, so Ward argues that all human belief systems involve myth-making. And this includes not just archetypal stories of our origins (such as the Genesis narrative), but a whole range of human activity —the 'symbolic realms we hominids have been cultivating for 2.2 million years', including art, poetry, rite, and dance. In ways we cannot fully explain, these interlocking modes of human culture tap the powers of what (for a want of a better term) we call the imagination, which operates at many more levels than are accessed by our conscious reflective awareness. Such works of the imagination 'intimate that our experience … of being in the world is freighted with a significance that only an appeal to the mythic can index' (Ward 2014: 186).

Many important issues are raised by this stress on the psychological depth and complexity of human cognition, and its mythical and the imaginative aspects. But for present purposes two key questions present themselves: first, what are the implications of all this for the epistemic status of religious belief; and second, what lessons emerge for the way in which the philosophy of religion should be conducted? To these questions we shall briefly turn in the fifth and final section of our discussion.

Philosophizing About Religion and the Layers of the Human Psyche

The line of argument canvassed in the previous section—emphasizing the creative, imaginative, and mythical elements in all human belief systems—might seem to offer a kind of protective armour for religious ways of thinking against the advances of modern scientific rationalism. But there may be grounds for concern that such a defence of religion is bought at too steep a cost—the cost of eroding the very distinction between *mythos* and *logos*, between imagination and reason. For even if science is necessarily the creation of our evolved human capacities, conditioned by our long human history, it has nevertheless developed tried and trusted methods (empirical investigation, mathematical modelling) for understanding and predicting the workings of nature. And the secularist charge against religious ways of thinking is that they completely fail to pass these tests for reliable belief formation, and thus do not deserve a place in our modern world view.

There is no space here to delve further into the extensive and continuing contemporary debate about the future of religion in the modern world. What needs to be addressed in the present context are the implications for philosophy, and in particular the philosophy of religion, of the issues raised in the previous section about the complex 'archaeology of belief'. In this connection, one does not have to sign up to a questionable assimilation of science and myth in order to wonder if contemporary analytic philosophy of religion has become too dry and austere, too closely modelled on the pared-down unambiguous language of the sciences, to do justice to the complexities of religious belief and the ways in which it might contribute to our understanding of ourselves and the nature of the reality of which we are a part. It is here that the contribution of psychoanalytic thought seems particularly relevant. For if there is any truth in what the psychoanalytic movement has tried to uncover about the hidden layers of the human mind, then it seems plausible to suppose that being more open to 'what lies beneath' might lead to a more nuanced epistemology, less modelled on the austere language and methods of the physical sciences, but arguably better equipped for the task of philosophizing insightfully about religion.[13] It is striking in this connection to find even a committed analytic philosopher of religion such as Eleonore Stump arguing recently that in order to do its job philosophy of religion may require deeper and richer resources than those afforded by the tools of logical analysis and technically expert argument (Stump 2010: 26–7).

Stump does not explicitly mention or invoke the resources of psychoanalytic theory, but significantly she does argue that philosophers of religion need to make use of our manifold responses to the multiple resonances of literary, and scriptural, narrative. This chimes in with earlier calls for a certain kind of narrative or literary turn in philosophy, as advocated for example in the work of Martha Nussbaum. Nussbaum argues that in learning to appreciate a great literary text we have to allow ourselves to be receptive and 'porous', knowing when to yield instead of maintaining constant critical detachment (1990: 281–2). Some philosophers may suppose that any departure from complete analytical detachment would involve a loss of philosophical integrity; and certainly there is need for philosophical caution whenever our imaginative and emotional resources are made use of. But equally, if we insist on maintaining a detached analytical stance at all times, this may be less a sign of intellectual integrity than what Nussbaum calls 'a stratagem of flight' (Nussbaum 1990: 268)—a refusal of the openness and receptivity that is prepared to acknowledge all the dimensions of our humanity.

If this is right, then one lesson to emerge is that we may need a new epistemology for thinking about religious belief and its basis. In contemporary analytic philosophy of religion, both the advocates of religious belief and its critics tend to operate with an *epistemology of control*. We stand back, scrutinize the evidence, retaining our power and autonomy, and pronounce on the existence or otherwise of God. But such methods implicitly presuppose that the divine presence ought to be detectable via intellectual analysis of formal arguments or observational data. Yet the ancient Judaeo-Christian idea of the *Deus absconditus* (the 'hidden God')[14] suggests a deity who is less interested in proving his existence or demonstrating his power than in the moral conversion and freely given love of his creatures, and in guiding the steps of those who 'seek him with all their heart', in Pascal's phrase (1670: no. 427). And when we start to think about the means of such conversion, it becomes clear that it could never operate through detached intellectual argument alone, or through the dispassionate evaluation of 'spectator evidence' (Moser 2008: 47).

Any suggestion that religious claims cannot fully and properly be evaluated from a detached and dispassionate standpoint may at first seem to be special pleading on behalf of religion; but further reflection makes it clear that there are all sorts of other areas of life—appreciation of poetry, of music, entering into any kind of personal relationship— where we need to be (to use Nussbaum's term) 'porous'. Otherwise, while we pride ourselves on being in control and judiciously evaluating the evidence, we may actually be closing ourselves off from allowing the evidence to become manifest to us. In short, there may be many areas of human life where a proper understanding of what is going on requires us to relinquish the epistemology of control and substitute an *epistemology of receptivity*.[15]

Perhaps surprisingly, this plea for the adoption of an epistemology of receptivity when assessing the claims of religion can draw some support from the writings of the founding father of psychoanalysis. Although as we have seen, Freud himself was a stern critic of religion, and although he tended to present himself very much in the garb of the austere

scientific rationalist, he also acknowledged, perhaps most explicitly in his *Introductory Lectures on Psychoanalysis*, that his methods of treatment diverged very significantly from those which are typical of standard scientific procedure. He points to two important differences. First, conventional scientific medicine looks to 'establish the functions and disturbances of the organism on an anatomical basis, to explain them in terms of chemistry and physics, and to regard them from a biological point of view', whereas Freud concedes that his own approach focuses on a much more elusive aspect, namely the 'psychological attitude of mind'. And second, Freud acknowledges the oddity of the fact that the processes involved are not susceptible of public investigation under normal observer conditions, because the psychotherapeutic process takes place in a private consulting room and 'only under the conditions of a special affective relationship to the physician' (1916–17: 22).

Thus, so far from there being an objective scientific template to which all valid discourse and all legitimate human inquiry must conform, Freud in the *Introductory Lectures* appears ready to allow that there are phenomena whose nature is such that quite different modes of understanding are appropriate. Indeed, he goes further and acknowledges that psychoanalysis is learnt first of all 'on oneself, by studying one's own personality' (1916–17: 23). These concessions are most significant, since (whether Freud himself drew such an inference or not) they implicitly cut the ground from underneath those critics of religion who would dismiss the validity of religious experiences on the grounds that they resist external verification by detached or non-involved observers.

The important lesson to emerge here is that despite the prevalence of scientistic modes of thinking in our contemporary culture (and in some parts of Freud's own thinking), we need to take seriously the idea that there may be phenomena that do not manifest themselves 'cold', as it were, but require involvement and commitment on the part of the subject in order to be apprehended. It has been an assumption of modern scientific inquiry that the truth is simply available for discovery, given sufficient ingenuity and the careful application of the appropriate techniques, and that the dispositions and moral character of the inquirer are entirely irrelevant. But while this assumption may be correct enough when inquiring into truths within meteorology, say, or chemical engineering, it seems quite out of place when we are dealing with certain central truths of our human experience—for example truths about how a poem or symphony may be appreciated, or how a loving relationship may be achieved and fostered. In these latter areas, the impartial application of a mechanical technique is precisely the wrong approach: the truth yields itself only to those who are already to some extent in a state of receptivity and trust.[16], [17] The upshot is that there may be phenomena, or parts of reality, whose detection or apprehension is subject to what might be called *accessibility conditions*: the requirements for getting in touch with them include certain requirements as to the subjective attitude and psychological (and perhaps moral) state of the subject (Cottingham 2005: ch. 1, §§ 3 and 4; ch. 7, §4). And as we have seen, Freud himself seems clearly to acknowledge this when he speaks of the insights arising from the psychotherapeutic process making themselves available 'only under the conditions of a special affective relationship to the physician' (1916–17: 23). Yet once it is granted that

there are psychological truths that may come to light only given certain affective and other transformations within the experiencing subject, then it may become possible to see how the same principle might be applied to religious truths, so that certain transformations in the subject may be crucially necessary preconditions for the manifestation of the divine reality that is the object of the religious quest (Cottingham 2009: ch. 5).

What thus emerges, as we bring to a close our discussion of the relation between psychoanalysis and religion, is the remarkable degree of convergence that obtains between these two very different ways of thinking about the human condition. Not only do both outlooks search for deeper layers of significance beneath of the surface world of factual assertion and plain 'common sense', but also, as we have just seen, both hold that this deeper world may disclose itself only to those who are in a suitable state of receptivity—a point that carries important epistemological implications perhaps not yet fully assimilated in our contemporary philosophical culture. To be sure, none of this is sufficient on its own to constitute a vindication of the claims either of psychoanalytic theory or of traditional religion, nor is it intended to be; but at least it may give some indication of the way in which those claims might have to be assessed.[18]

References

Augustine of Hippo (*c.* 398 **CE** [1912]). *Confessions* (*Confessiones*), trans. W. Watts. Cambridge, MA: Harvard University Press.

Cottingham, J. (1998). *Philosophy and the Good Life: Reason and the Passions in Greek, Cartesian and Psychoanalytic Ethics*. Cambridge: Cambridge University Press.

Cottingham, J. (2005). *The Spiritual Dimension*. Cambridge: Cambridge University Press.

Cottingham, J. (2009). *Why Believe?* London: Continuum.

Cottingham, J. (2014). *Philosophy of Religion: Towards a More Humane Approach*. Cambridge: Cambridge University Press.

Cottingham, J. (2015). *How to Believe*. London: Bloomsbury.

Dante Alighieri (*c.* 1310 [1981]). *The Divine Comedy: Paradise* (*La Divina Comedia: Paradiso*), ed. and trans. G. Bickersteth. Oxford: Blackwell.

Dawkins, R. (2006). *The God Delusion*. London: Transworld Publishers.

Foucault, M. (2005). *The Hermeneutics of the Subject*. New York: Palgrave Macmillan.

Freud, S. (1901). *The Psychopathology of Everyday Life*. In J. Strachey (ed. and trans.), *The Penguin Freud Library*, Vol. 5. London: Penguin.

Freud, S. (1916–17 [1952]). *Introductory Lectures on Psychoanalysis* (*Vorlesungen zur Einführung in die Psychoanalyse,*), trans. J. Riviere as *A General Introduction to Psychoanalysis*. New York: Washington Square Press.

Freud, S. (1927). 'The Future of an Illusion'. In J. Strachey (ed. and trans.), *The Penguin Freud Library*, Vol. 12 (pp. 182–241). London: Penguin.

Freud, S. (1929). 'Civilization and its Discontents'. In J. Strachey (ed. and trans.), *The Penguin Freud Library*, Vol. 12 (pp. 251–340). London: Penguin.

Howard-Snyder, D. and Moser, P. (eds) (2002). *Divine Hiddenness*. Cambridge: Cambridge University Press.

Hume, D. (1757). 'The Natural History of Religion'. In *Four Dissertations*. London: Millar.

Jung, C. (1918). 'The Role of the Unconscious' ('Über das Unbewußte'). In C. G. Jung, *Collected Works*, rev. edn (l967–77), Vol. 10. London: Routledge.

Jung, C. (1938). *Psychology and Religion*. In C. G. Jung, *Collected Works*, rev. edn (l967–77), Vol. 11 (pp. 89–95). London: Routledge.

Jung, C. (1939). 'Conscious, Unconscious and Individuation'. In C. G. Jung, *Collected Works*, rev. edn (l967–77), Vol. 9 (1) (pp. 275–289). London: Routledge.

Jung, C. (1951). *Aion*. In C. G. Jung, *Collected Works*, rev. edn (l967–77), Vol. 9 (2). London: Routledge.

Jung, C. (1952). 'Symbols of Transformation'. In C. G. Jung, *Collected Works*, rev. edn (l967–77), Vol. 5. London: Routledge.

Jung, C. (1956). 'Correspondence with H. L. Philp'. In C. G. Jung, *Collected Works*, rev. edn (l967–77), Vol. 18 (pp. 706–707). London: Routledge.

Kahneman, D. (2011). *Thinking Fast and Slow*. London: Penguin.

Kihlstrom, J. (1987). 'The cognitive unconscious'. *Science 237*: 1445–1452.

Knox, J. (2003). *Archetype, Attachment, Analysis*. Hove: Brunner-Routledge.

Leiter, B. (2004). *The Future for Philosophy*. Oxford: Clarendon Press.

Martin, T. (2002). *Oppression and the Human Condition*. Lanham, MD: Rowman & Littlefield.

May, S. (2011). *Love: A History*. New Haven, CT: Yale University Press.

McGilchrist, I. (2009). *The Master and His Emissary*. New Haven, CT: Yale University Press.

Moser, P. (2008). *The Elusive God: Reorienting Religious Epistemology*. Cambridge: Cambridge University Press.

Nagel, T. (1974 [1979]). 'What is it like to be a bat?'. In *Mortal Questions* (ch. 12). Cambridge: Cambridge University Press.

Nagel, T (1994). 'Freud's Permanent Revolution'. *New York Review of Books 41*(9): 34–38.

Nielsen, J. et al. (2013). '**An Evaluation of the Left-Brain vs. Right-Brain Hypothesis**'. *Plos* One *8*(8): e71275.

Nussbaum, N. (1990). *Love's Knowledge*. Oxford: Oxford University Press.

Palmer, M. (1997). *Freud and Jung on Religion*. London: Routledge.

Pascal, B. (1670 [1962]). *Pensées*, ed. L. Lafuma. Paris: Seuil.

Ricoeur, P. (1970). *Freud and Philosophy: an essay on interpretation*, trans. D. Savage. New Haven, CT: Yale University Press (Originally published as *De l'interprétation: Essai sur Sigmund Freud*, 1965).

Sartre, J.-P. (1975). *The Emotions: Outline of a Theory*, trans. B. Frechtmann. Secaucus, NJ: Citadel Press (Originally published as *Esquisse d'une théorie des emotions*, 1939).

Stump, E. (2010). *Wandering in Darkness*. Oxford: Oxford University Press.

Ward, G. (2014). *Unbelievable: Why We Believe and Why We Don't*. London: I. B. Tauris.

Notes:

(1.) Quoted in Martin 2002: 67. It should be noted that Sartre himself was critical of the Freudian concept of the unconscious mind.

(2.) [Eds: See Blass, this volume, § 'The Dialogue Between Freud and the Believer on the Nature of Truth', for an account of Freud's later explanation, in *Moses and Monotheism*, of the force of religious belief in terms of 'historical truth'.]

(3.) [Eds: Cottingham's argument may be fruitfully juxtaposed with Blass, this volume, § 'Contemporary Positive Perspectives' and Boothby, this volume.]

(4.) 'Fecisti nos ad te, et inquietum est cor nostrum donec requiescat in te' (Augustine (*c.* 398), Bk. I, Ch. 1; 'E'n la sua volontade è nostra pace' (Dante Alighieri (*c.* 1310), Canto iii, line 85.

(5.) [Eds. See also Blass, this volume, § 'Contemporary Positive Perspectives'.]

(6.) For more on the psychodynamics of this transformational process, see Cottingham (1998), ch. 4.

(7.) 'Hermeneutics' in the most general sense may be thought of as an approach to philosophy which gives a central place to the (culturally mediated) search for self-understanding. See Ricoeur 1970.

(8.) [Eds. Blass, this volume, § 'The Change in the Attitude to Truth' explores Freud's antipathy to this turn in relation to questions of religion.]

(9.) [Eds. See Gipps, this volume, for a development of this line of thought.]

(10.) See McGilchrist, 'Exchange of Views' at http://iainmcgilchrist.com/exchange-of-views/.

(11.) See for instance Eleonore Stump on 'cognitive hemianopia' (2010: 26–7).

(12.) Compare McGilchrist 2009: 8.

(13.) [Eds. A similar conclusion is reached via a very different route by Blass, this volume.]

(14.) See Isaiah 45:15. For more on the 'hiddenness' of God, see Howard-Snyder and Moser (2002).

(15.) The argument of the last three paragraphs draws on material from Cottingham 2014: ch. 1. See also Cottingham 2015, esp. ch. 3.

(16.) For more on this theme, see Michel Foucault, Seminar at the Collège de France of 6 January 1982, trans. in Foucault (2005: 1–19). See also Cottingham 2005: ch. 5 and ch. 7.

(17.) [Eds. For development of this line of thought in relation to psychoanalysis, see Gipps, this volume.]

(18.) I am grateful to Richard Gipps and Michael Lacewing for their most thoughtful comments on an earlier draft of this chapter.

John Cottingham

John Cottingham is Professor Emeritus of Philosophy at the University of Reading, UK.

Access is brought to you by

Oxford Handbooks Online

The No-Thing of God: Psychoanalysis of Religion After Lacan 🔒

Richard Boothby

The Oxford Handbook of Philosophy and Psychoanalysis
Edited by Richard G.T. Gipps and Michael Lacewing

Subject: Philosophy, Philosophy of Mind, Philosophy of Religion
Online Publication Date: Nov 2018   DOI: 10.1093/oxfordhb/9780198789703.013.35

Abstract and Keywords

This chapter traces some main lines of Jacques Lacan's interpretation of religion and divinity, which differs significantly from Freud's critique. Orienting ourselves with respect to what Lacan calls das Ding, the enigmatic desire of the Other, it is possible to sketch a Lacanian analysis of religion parallel to that offered by Kant in Religion Within the Limits of Reason Alone. The difference is that where Kant looked to find in religious representations, especially those of Christianity, the underlying dynamics of pure rationality and of a morality founded upon it, Lacan discerns the very structure of subjectivity and its relation to the unknown Other. New perspectives are thereby opened up on a whole series of problems, including the unconscious dynamics of enjoyment, practices of sacrifice, the structural differences between various religions, and Christian doctrines of incarnation, love, and mystical unknowing.

Keywords: Lacan, Real, Imaginary, Symbolic, the Other, das Ding, paternal metaphor, sacrifice, sexuation, mysticism

Introduction

On the topic of religion, it is clear that the French psychoanalyst, Jacques Lacan, doesn't merely repeat the critique mounted by Freud. Not so clear is the meaning of what Lacan says on his own.[1] At different points in his teaching, Lacan associates the ancient gods with the unthinkable Real (2017: 44), he identifies the 'I am that I am' of the burning bush with the pure subject of speech (1968: Session of 11 December), he asserts that 'the true formula for atheism is that God is unconscious' (1981: 59), and he intimates that his *Écrits* should be compared with the writings of the great mystics (1998: 76). That last provocation is sounded in Lacan's twentieth seminar, *Encore: On Feminine Sexuality, The Limits of Love and Knowledge*. It is there, too, that Lacan associates an excessive enjoyment beyond knowing, what he calls the '*jouissance* of the Other', with the essential incompleteness of woman, her characteristic of being *pas tout*, or 'non-all'. He then brings such feminine *jouissance* suggestively but enigmatically into conjunction with God, asking at one point: 'Why not interpret one face of the Other, the God face, as based on feminine *jouissance*?' (1998: 82).

How to make sense of these obscure passages? How are we to grasp the intersection they imply between *jouissance*, language, lack, love, and God? And how are these terms knotted in relation to knowing and what is beyond knowing? While such questions cannot be adequately answered in a brief paper, we might hope to provide a basic orientation to them.

Freud's Critique

Freud's critique of religious belief is familiar and need only be tersely summarized to orient our discussion.[2] In *The Future of an Illusion* (1927), Freud's attack is directed less towards the *object* of belief, the dubiousness of an invisible, supernatural power who oversees human life from a cosmic distance, than towards the *subject* who believes. The problem is the motive of the believer. The 'illusion' at stake is thus a matter of the believer's investment in wish-fulfilling fantasies.

To this bottom-line assessment, Freud added two more properly psychoanalytic angles of interpretation. The first, developed in *Totem and Taboo* (1911) and revisited in Freud's last major work, *Moses and Monotheism* (1939), predicates the solidarity of the human group upon the fantasy of the father's murder. The bond of brotherly love is cemented by the blood of the patriarch. It is a theory that casts a light on the origins of religious ritual in practices of blood sacrifice and suggests that what Catholics call the mystical body of the Church is founded upon a collective guilt over the murder of the Saviour.

Freud based a second strategy on the resemblance of religious ritual to the compulsions of the obsessive neurosis. Not only are they both marked by a conspicuous conscientiousness, they also display the compromise structure of the symptom, which functions to repress some impulse at the same time that it symbolically discharges it. The case of the Ratman illustrates the dynamic perfectly (Freud: 1909: 153–320). His obsessional thinking becomes unmanageable upon hearing a fellow soldier describe an oriental torture in which a cage of starving rats is affixed to the buttocks of the victim so that the rats chew their way into the anus. The Ratman becomes preoccupied with the recurring dread that such horrifying treatment might be applied to the two most important figures in his life: his father and his fiancée. In the frightful shudder that accompanies each compulsive reimagining, Freud recognizes both a struggle to defend against a sexually charged desire and a means of satisfying that very desire. Key religious phenomena might be interpreted in parallel fashion. Contemplation of the crucifixion, for example, invites the worshipper to rise above all hatred in a final victory of love while also offering an occasion to gratify an unconscious sadism.

Lacanian Heresy

In a simple but elegant pun, the title of Lacan's twenty-second seminar, 'R.S.I.,' ––which pronounced in French sounds like 'heresy'––suggests both a measured departure from Freudian orthodoxy and, more subtly but palpably, a promise of a new angle of interpretation on the problem of religion. The three letters stand for the three fundamental registers––Real, Symbolic, and Imaginary–– with which Lacan rewrites the essentials of psychoanalytic theory. The case of the Ratman offers a useful frame in which to introduce them.

The fantasy of bodily violation in the rat torture recalls the dynamics of the 'imaginary', which for Lacan inform the structure and function of the ego. During the period of late infancy that Lacan calls 'the mirror stage', the unifying gestalt of the other's body is thought to coordinate the impulse chaos of the child. This reliance on an external image alienates the human subject from itself. It inevitably invites aggression towards the other because it introduces an internal split. On this point, Lacan rests a central postulate of his entire outlook: the opposition between the *moi* and the *je*, between the *ego* as the seat of one's conscious sense of self, and the *subject* of unconscious desire.

That subject emerges in pulsations of the 'symbolic', Lacan's general term for the resources of language, informed by the theories of Saussure and Jakobson, on which he draws in order to retrieve and radicalize Freud's many analyses of word play in dreams, slips, bungled actions, and symptomatic complexes. In the Ratman, for example, the obsessive idea clearly crystallizes around the morpheme 'Rat', which functions like a kind of switch-point in a constellation of emotionally loaded issues: the starving, gnawing rodents (*Ratten*), monetary debts (*Raten*), gambling (*Spielratte*), and marriage (*Heiraten*). The key thing that Lacan adds to Freud's own elaborate analyses (and what he adds to the linguistic theories of Saussure as well) is a more general theory of the signifier centred on its unconscious dimension. The signifier always passes into the speech stream an indeterminate excess, what Lacan calls the '*objet petit a*'. In the margin of this excess there is a 'slippage of the signified beneath the signifier' (Lacan, 2006: 419) The consequence, put in simplest terms, is that we always say more than we intend. It is in the space of the unsaid and unintended that unconscious effects unfold.

By locating the roots of unconscious desire in the play of the signifier, Lacan reconceives the Oedipus complex as a function of the acquisition of language.[3] What Freud had tied to paternal threats of castration becomes a matter of tensions internal to the child's submission to the symbolic law. Lacan's third category of the 'real' arises at the limit of the law, or better, at the point where the symbolic function remains incomplete, interrupted, or inconsistent. In the gaps and failures of the symbolic, the subject may be confronted with a traumatic excess, the subjective impact of the real that Lacan calls *jouissance*. The real in this sense, far from being reducible to everyday 'reality', indicates the subject's encounter with something fundamentally unthinkable. If it is an encounter with something incomprehensible in the outside world, it is also an engagement with the

incomprehensible in the subject itself. Here, too, the Ratman provides us with a suggestive example. In talking about the dread rat torture, Freud tells us, the patient's face 'took on a very strange, composite expression. I could only interpret it as one of *horror at pleasure of his own of which he himself was unaware*' (Freud 1909: 166–7).

Where, then, are we to locate the religious in the architecture of Lacanian theory? For many commentators, the answer is to identify God with an aspect of the unthinkable real. Divinity is to be situated in a domain beyond human experience, a zone that forever outstrips all capacities to image or name it.[4] Such an approach offers the satisfaction of locating the divine in a dimension of radical transcendence, but it also runs the risks attendant upon the conception first put forward by Kant, who famously claimed to limit human reason in order to make room for God. For such a view, the divine is vouchsafed by being located in a noumenal Beyond, but only at the price of our being barred from saying anything whatever about its nature.

A second angle of interpretation takes its point of departure from Lacan's conception of the symbolic 'big Other'. Part of what is distinctive about Lacan's conception of language is that he characterizes the network of the symbolic code as susceptible of being construed as something akin to, or in some way inhabited by, another subject. This so-called 'big Other' is the nameless and faceless regulator who oversees the written and unwritten rules that direct our lives. At stake is not just the properly linguistic rules governing grammar, syntax, and semantics. The big Other also monitors our polite behaviour, inspects our adherence to fashion, defines for us what is funny (sometimes actually laughing for us in the 'laugh track' of TV sitcoms), insists on our silent decorum at one moment and calls for our deep-throated patriotic fervour at another. The power of this symbolic big Other resides in part precisely in the fact that it exists everywhere and nowhere, that it is both 'all and none'.[5]

In this supposition of something subject-like in the symbolic code we rightly discern a shadow of Freud's primal father. Yet we also get a hint that, unlike Freud, who posed faith as a means of coping with human helplessness by means of pleasing illusions, for Lacan the disposition to belief in God is more emphatically unconscious and more definitely structural. Far from being an illusory add-on to a rationality that could and should do without it, an implicit 'faith' in the underlying coherence of sense-making is a kind of structural requirement of our relation to language. The Lacanian view thus suggests that something like faith, the baseline credulity that undergirds reliance on all symbolic conventions, is a virtual necessity for human beings. It is this function that underlies the promotion of what Lacan calls 'master signifiers', the legitimacy of which are warranted by the big Other in the role of a 'subject supposed to know'. And what is God but the ultimate form of such supposition? If we mortals fail to grasp the meaning of our lives, the same is not true of the divine agency whose plan for Creation is complete in every detail.[6] 'The subject supposed to know', says Lacan, 'is God, full stop, nothing else' (1968: Session of 30 April).

God in the real or in the symbolic. Must we make a choice? A prime purpose of the remainder of this chapter is to show how the two perspectives are in fact inseparable. The key for making their connection lies in Lacan's concept of *das Ding*, the Thing.

In the Shadow of the Thing

While mentioned only infrequently after the central role it plays in the Lacan's seventh Seminar on '*The Ethics of Psychoanalysis*', the concept of the Thing is clearly fundamental.[7] As Lacan says of it, '*Das Ding* is a primordial function which is located at the level of the initial establishment of the gravitation of the unconscious *Vorstellungen*' (1992: 62). How, then, to interpret it? Much of the commentary on Lacan's notion has followed a Kantian clue. On this account, the Lacanian Thing is akin to the Kantian *Ding-an-sich*, the unthinkable kernel of external things, not merely as they appear to us, but as they are in themselves. It is along this line of thinking that Lacan's Thing is associated with the original lost object that sets in motion the ceaseless roving of human desire. *Das Ding* also names the dimension of what is fundamentally excessive (and thus is connected in Lacan's thought to the *object petit a*). As such, it animates the experience of the sublime, which Lacan famously characterizes as emerging when 'the object is elevated to the dignity of the Thing' (1992: 112).

This quasi-Kantian posing of the Thing in the locus of the *Ding-as-sich* is by no means without some value, but it needs to be resisted at least long enough to recognize that the crucial dimension of the *das Ding* concerns not objects but other people. The original unthinkable object is the fellow human being. This conclusion is obvious when we return to the text of Freud from which Lacan takes his clue. In a brief flight of theorizing in his unpublished 'Project for a Scientific Psychology', Freud notes how the child divides the figure of the *Nebenmensch*, the 'neighbour' or fellow human being, between what it can recognize on the basis of similarities to its own body––precisely the sort of mirror recognition that Lacan associates with the imaginary—and a locus of something that is 'new and non-comparable', a zone of something unknown. This unrepresentable excess Freud calls *das Ding*. It is this division of the human other, a division that reserves in the heart of the familiar a locus for something excessive and as yet unknown, that will serve as the template for all of the child's future attempts to interrogate the nature of objects. 'For this reason', says Freud, 'it is in relation to the fellow human-being that a human-being learns to cognize' (1886–99: 331) The key point is not merely that there remains a inaccessible, noumenal core of all objects but that the original schema that locates such an unknown is taken from the example of what is unknown in the human Other.[8]

What is new and crucially important in Lacan's treatment of the Thing is the way in which the enigmatic locus of something uncognized in the Other becomes the root source of anxiety. 'Not only is [anxiety] not without object', he says, 'but it very likely designates the most, as it were, profound object, the ultimate object, the Thing' (Lacan 2014: 331). The challenge of the neighbour-Thing consists not simply in the discovery of an inaccessible kernel at the heart of the Other but in the way it raises the unsettling question of what object I am for that unknown desire. The question presses with particular force in the drama of toilet training when the demand of the Other for the regulation of the infant's bowels re-energizes anxiety about the unanswered question of

what the Other wants. The Lacanian thesis thus goes beyond merely locating the source of anxiety in the fellow human being. It asserts, contrary to our fondest myth about childhood, that the hidden source of deepest and most uncanny anxiety is the mother herself.[9]

We might remark in passing how this Lacanian understanding of the dynamics of the mother/child relation at least partly resembles that offered by Simone de Beauvoir. At a key point of her argument in *The Second Sex*, de Beauvoir reasserts the value of the Freudian Oedipus complex for understanding the deep roots of ambivalence towards the feminine, though with the crucial proviso that we invert Freud's conception (de Beauvoir 1989: 195–6). The core of the Oedipus complex is not, as Freud thought, that the child must be separated from the mother by a threat of castration, but rather that the child is motivated to navigate its own separation, seeking to achieve an autonomy that can be won only by a certain rejection of the maternal embrace. Lacan echoes this key point. He could well be paraphrasing de Beauvoir when he insists that 'it's not true that the child is weaned. He weans himself. He detaches himself from the breast' (Lacan 2014: 313).

Does Lacan then follow de Beauvoir here? By no means completely. Their point of convergence only makes it more essential to clarify Lacan's distance from de Beauvoir, who departs from Freud merely in claiming that the child's desire for the mother is abrogated not by the threat of paternal castration but rather by the child's own rejection. For Lacan, by contrast, the problem isn't the desire of the child at all, but rather that of the mother, in as much as her desire is encountered as a threatening unknown. It is in the light of this perspective that we can make sense of Lacan's comparison of the mother to the daunting spectre of a giant praying mantis (Lacan 2014: 22). In the same stroke, we can interpret his characterization of the *objet petit a* as *un objet cessible*, a cedable or yieldable object (see Lacan 2014: 312–15, 324–7). In the various incarnations of the *objet a*, most of them parts of the body, the subject's 'pound of flesh', is exchanged or 'sacrificed' in order to create a margin of safe separation from the Other.[10] Lacan calls to mind a similar function when he likens the mother to a crocodile and the phallic signifier to a stick with which to keep its jaws from snapping shut.

The notion of the *objet cessible* recapitulates in Lacan's own theoretical frame something akin to Heidegger's notion of an escape from anxiety made possible by means of focusing on an object (Heidegger 1962: 234). What Lacan calls *objet a* ultimately derives from a pure void, a pure function of lack and absence. Yet this lack is successively given definite shape in a series of objects, in the first instance parts of the body (breast, faeces, penis etc.).[11] In this way, the yawning lack discovered in the unfathomable opacity of the Other our topic, the result is a new angle of view on the meaning of sacrifice. In its ultimate function, 'sacrifice is not at all intended to be an offering, nor a gift, both of which are propagated in a quite different dimension, but the capture of the Other in the web of desire' (2014: 277).

This discussion of the materialization of the *objet a* and its deployment in an economy of exchange provides a ready segue to the child's entry into language, the most basic and highly elaborated economy of exchange. In the originary drama with the maternal Other, the inarticulate cry of the infant becomes in itself a ceded object, indeed the very first such object, yielded into the space between subject and Other. As Lacan says, 'this manifestation of anxiety coincides with the very emergence in the world of he who is going to be the subject. This manifestation is his cry … this first effect of cession … the nursling can't do anything about the cry that slips out of him. He has yielded something and nothing will ever conjoin him to it again' (2014: 326).

If the first inchoate eruptions of the voice are in this way inflected with anxiety, and inevitably so, in as much as they enter and begin to shape the emerging interval between the subject and the Other, they are also destined to become the means by which the question of the Other, the enigma of *das Ding*, will be ceaselessly reposed. 'Here we're touching', says Lacan, 'on the very thing that makes the relation to the Other possible, that is, on that whence emerges the fact that there is such a thing as a signifier. This site whence emerges the fact that there is such a thing as the signifier is, in one sense, the site that cannot be signified. It is what I call the site of the lack-of-signifier' (2014: 134). What is at stake is the most elemental implication of Lacan's theory of the signifier, perhaps precisely what he had in mind by claiming that he had defined the signifier as no one else had dared. In every entry into language, in every iteration of signifying material, there resounds some echo of the unanswered and unanswerable question of the Other.

This perspective on *das Ding*, in which the primordial question of the Other resonates–– waveringly, uncertainly––every signifier, illuminates the point that Lacan never tires of emphasizing: that it is always possible to ask about anything spoken, 'yes, I hear what you're telling me, but what are you really saying (what do you really want)?'. What most distinguishes Lacan's view of language and its function is that meaning can never be fully stabilized, that a question not only can but always implicitly *is* posed by every entry into language. Yet if every play of the signifier implicitly resumes an interrogation of the unknown in the Other, what is ultimately at stake is the subject's own coming-to-be, the subject's own question. This even more elusive stake of the game is rooted in the real of the subject's mute *jouissance*. It is a locus that can be approached only mythically, conceivable only as a primordial, indeterminate X. The key point for Lacan, however, is that the question of the Other always comes first, that the question of the subject's desire can be posed only in the locus of the Other.[12] 'What anxiety targets in the real'. Lacan says, 'includes the *x* of a primordial subject moving towards his advent as subject, […] since the subject has to realize himself on the path to the Other. […] [this subject] is the subject of *jouissance* […] it can in no way be isolated as a subject, unless mythically' (2014: 173).

Hell is Other People

The upshot of the foregoing discussion, as shockingly violent to common sense as it is to mainstream linguistics, is to assert that the most archaic function of language, far from connecting the subject to the fellow human being, is to achieve an indispensable degree of separation from the Other, to establish a margin of detachment and independence that puts the neighbour-Thing at a distance.[13] The primordial accomplishment of the signifier is to provide a stand-in for the unknown dimension of the Other, putting that threatening unknown at a remove and thereby forestalling the outbreak of anxiety.

With this process in view, we can say with fresh literalness that 'the word is the murder of the Thing'. The phrase refers to the philosophy of Hegel.[14] Yet if the Hegelian formula now becomes meaningful in a new way, another Hegelian concept becomes even more relevant, insofar as the function of the signifier becomes an exemplary instance, indeed *the* exemplary instance of Hegel's cardinal concept of *Aufhebung*, which forms the inner pivot of his famous dialectic.[15] In accord with the two-fold meaning of the German word, something that undergoes *Aufhebung* is simultaneously *cancelled and preserved*. In exactly such a dialectical movement, the signifier both cancels *das Ding*, distancing the subject from it, offering an arbitrary sound in the place of the unknown Thing, yet also preserves precisely what is cancelled, marking it for further coginizing sometime in the future. The signifier establishes a locus suspended between the subject and the Other, at once putting what remains unknown and potentially threatening about the Other at a safe distance, yet thereby repositioning it and maintaining it as a question.

This perspective integrates the most basic points of Lacan's theoretical framework. To identify the elementary function of the signifier with an *Aufhebung* of the enigmatic Thing in the Other is nothing other than to repose the basic terms of Lacan's 'paternal metaphor', his way of reconceptualizing the Oedipus complex, in which a signifier, what Lacan calls the 'Name of the Father', is substituted for the unknown Desire of the Mother. In the same stroke, it explains how and why the inauguration of the signifier is already and in itself the most primitive inscription of the law against incest. There can be no subject without opening a certain distance from the maternal Thing.

To summarize the view we're adopting here, the very being of the subject is predicated upon a suspension of the unknown reality of the Other, the continued management of which becomes the charge of the symbolic law. This perspective is audible in a question put forward by Slavoj Žižek, apropos Lacan's distance from Levinas: 'What if the ultimate function of the Law', Žižek asks, 'is not to enable us not to forget the neighbor, to retain our proximity to the neighbor, but, on the contrary, to keep the neighbor at a proper distance, to serve as a kind of protective wall against the monstrosity of the neighbor? (2005: 163). From this perspective, the lack in the Other, the enigma of the Other's

jouissance, becomes the undetectable 'dark matter', by definition invisible to ordinary consciousness, that underlies the entire edifice of the social world.

We can now return to the question of religion and God with new resources. I'll limit myself to sketching four brief points, by necessity more suggestive and programmatic than comprehensive.

Economies of Jouissance

The analysis we have pursued allows us to trace the basic outline of a Lacanian interpretation of the religious in a manner reminiscent of Kant's (1793) effort in *Religion Within the Limits of Reason Alone*, highlighting those features of religious practice and belief that dovetail with features of human life discernible by wholly secular analysis. The difference is that where Kant's approach was oriented towards the human capacity for reason, for Lacan the crucial problem is the very structure of subjectivity, a structure, as we have seen, fundamentally oriented towards the desire of the Other. Where Kant remains in the orbit of *cogito*, Lacan is attuned to *desidero*. The crucial dimension is not thought but desire, not rationality but *jouissance*.

In Seminar 16––the seminar entitled '*D'un Autre à l'autre*' ('From an Other to the other') —Lacan offers a particularly pithy note about how the big and little Others are to be distinguished, a note that reminds us that the problem of the unknown in the Other is always a question of an unknown enjoyment. The little Other, the fellow human being or neighbour, is the locus of an incipient *jouissance*, the big Other is a second locus that has been somehow cleared, or evacuated, of enjoyment.

> This neighbor, is it what I have called the [big] Other, what I make use of to make function the presence of the signifying articulation in the unconscious? Certainly not. The neighbor is the intolerable imminence of enjoyment. The [big] Other is only its cleared out *terreplein*. […] It is precisely that, it is a terrain cleared of enjoyment

> (1968: Session of 12 March).

This passage points to an elemental shift in which an oppressive and intrusive presence of enjoyment in the little Other is traded for some kind of 'clearing', or leveling of enjoyment in the big Other. The implication is that the move towards a reliance on a big Other of the symbolic code answers to the challenge of the neighbour-Thing by means of 'taming' the desire of the Other. We're invited to see the so-called 'clearing' of enjoyment as a structural moment, a kind of necessary base camp or staging, on the way to the establishment of a regulated economy of enjoyment. What sort of regulation is this and how does it work?

The key to moving forward from this point and making new sense of the question of religion and God is to recognize that Lacan's reference to a 'clearing' or evacuating of enjoyment in the big Other doesn't mean that the locus of the big Other remains forever empty of *jouissance*. On the contrary, Lacan refers repeatedly to the big Other as a special site of projected enjoyment. The big Other comes to be another locus, another scene, of enjoyment but one that bears a special relation to the structure of the Law. From this vantage point, religious belief presents us with the spectacle of a congregation of human beings, gathered in the common cause of controlling too raw an exposure to one another's enjoyment in favour of devotion to another, imagined subjectivity in the form of a deity who controls access to enjoyment in accordance with sacred dictates.[16]

Such an economy of *jouissance* is shown with particular clarity, and with particular relevance to questions of the religious, in Lacan's conception of masochism and sadism. His primary point is to correct the vulgar conception of the masochist as someone who simply enjoys pain and the sadist as one who enjoys inflicting it. What is crucial, Lacan insists, is the way in which both masochism and sadism are triangulated by reference to the locus of the big Other. The masochist submits to pain in service of the larger purpose of inciting the anxiety of the big Other. The sadist, for his part, seeks to make himself the instrument of the big Other's enjoyment.

This rereading clarifies a great deal, not least the theatricality of masochism and sadism, the tendency to elaborately *stage* torments. The ultimate spectator of perversion is the gaze of the big Other. Among other things, this means that the results of the perverse game for the little Others—the perverse couple who enact it—are typically not only different but virtually the opposite of what they are for the big Other. In the case of the masochist, far from inciting the anxiety of the torturer, the masochist's submission offers to the sadistic partner a thrill of power, a complete escape from anxiety (the beating scenes from Lars Von Triers' film *Nymphomaniac* illustrate the dynamic perfectly). The anxiety is produced not in the human Other, but wholly on the level of the big Other. It is the big Other whose rules are violated, who is scandalized by pain, injury, and death. Likewise, in sadism, the crucial enjoyment is not that of the torturer (who, as we see in Sade's endless scenarios of torment, is virtually mechanical, unfeeling) but the big Other, who enjoys in his place. In the tenth Seminar, Lacan explicitly makes the link to the religious question: masochism and sadism are devoted respectively to the anxiety and the *jouissance* of God (Lacan 2014: 163). The conclusion to be drawn would have been familiar to St Augustine. Insofar as perversion is always played out under the gaze of the big Other, the perverse is coincident with the dimension of the religious.

From Mortals to Divinities ... and Back Again

For Lacan, the initial separation of the child from the mother is achieved not under pressure of a paternal threat of castration, but by means of the emerging agency of language. The relevance of this point for a new psychoanalytic interpretation of religion is compactly signalled by the opening line of John's Gospel: 'In the beginning was the Word'.

It is now possible to expand upon that mapping of the primal origin of the subject by emphasizing the role of the signifier as an *objet cessible*, a 'ceded object' injected into the space between the subject and the Other. This mediating object can be related to what is arguably the most elemental gesture of all religious life: the act of sacrifice. The signifier is itself the primal instance of something given up by the subject in order to found an economy of exchange with the Other. The sacrificial offerings of religion are essentially signifiers whose function is to establish the field of mediation between subject and Other. [17]

As we have seen, that process of signification inevitably tends towards unfolding of a third locus, that of the symbolic big Other, that grounds the consistency of the code, thereby stabilizing the subject's relation with the little Other of the fellow human being. It is with reference to this triangular structure of the subject, the little Other, and the big Other that Lacan is able to explain the source and strength of the disposition to belief in gods. Where Freud starkly opposed the blindness of faith to all rationality, Lacan roots the motive sources of belief in the elemental architecture of subjectivity. At the most fundamental level, the appeal to deities is grounded in the necessity for stabilizing the referential valence of linguistic signifiers. [18]

Reference to the elementary triangle of subject, little Other, and big Other also promises a means of explaining the logic of a certain evolution of religions, a sort of phylogenesis of religious life, moving across the trajectory from polytheism to monotheism and displaying a marked tendency towards collapsing the hypotenuse of the triangle that separates the little and big Others.

At one extreme, we have something like Aztec sacrifice, centred on an absolutely inhuman Other (the sun) that requires unending human sacrifice. The divine being is a sort of fearsome cosmic parasite that feeds ceaselessly on spurting human blood.

In ancient polytheistic paganism, the gods still require copious sacrifice though a new, much more advantageous arrangement is arrived at in which, as Hesiod tells us, humans get to eat the slaughtered victims and the gods content themselves with the smell of roasting meat wafting towards the heavens. In pagan sacrifice, specific objects are exchanged between mortals and divinities, stabilizing a link between them but also maintaining an unbridgeable distance.

In Judaism, the deity becomes a single individual and the dynamic of sacrifice alters radically. To seal the covenant with Abraham, Yahweh wants only the foreskin, thus reducing the stakes of sacrifice to an almost purely symbolic exchange. The focus on such a relatively trivial object, a conspicuously expendable part of the body, albeit one attached to the privileged organ of vitality, indicates that the value of sacrifice is no longer to be measured by the number or value of victims offered but entirely by the willingness of the sacrificer to give it up. The upshot is to signal an interest less in the

object of sacrifice than in the inward intention, the allegiance and obeisance, of the one who sacrifices.

To this basic view, the commentary of Slavoj Žižek adds an interesting sidelight. Žižek argues that the Hebrew prohibition of idolatry should be interpreted as a response to the fact that the deity has become so close to being human that there arises a kind of anxiety of coincidence, a fear, in effect, that the god is too close for comfort. In the terms we've been forging, this proximity threatens to collapse of the distance between the little and big Others, in turn undermining the economy of enjoyment that is predicated upon it. The problem with graven images and idols is not that they fail to do justice to the transcendent majesty of God but rather, as Žižek puts it that they 'would render [the Jewish God] too faithfully, as the ultimate Neighbor-Thing' (2001: 130–1).

Yet it is precisely such an identification of the god with the neighbour-Thing that is proposed by Christianity. The Christian 'fulfilment of the law' is achieved through a radical reversal of the order of sacrifice, a reversal that is, as Paul says, folly for the Greeks and scandal for the Jews: God sacrifices to save man. Even more radically, God becomes man. The Christian incarnation thus completely collapses the distance separating the big Other and the little Other. The new God can now say to his disciples: 'As you do to the least among you, you do unto me'. 'I am there wherever two or more of you are gathered'.

Sexuation and the Sacred

If the function of the religious centrally involves establishing an economy of *jouissance* that is triangulated between the little and big Others, then we can expect the essential structures and internal tensions of religious life to reflect Lacan's theory of sexuation, by which he sought to distinguish masculine and feminine forms of *jouissance*.[19] To see how they do, we need a little more background.

Among the most important and influential features of Lacan's late teaching are his so-called 'formulae of sexuation', his gloss on Freud's famous question about 'what does a woman want?'. His scheme derives from a deliberate torsion––many will say simply 'distortion'–– of classical logic, resulting in two different modes of relating to the universal. The first, characteristic of the masculine position, defines the universal by reference to the exception. As Freud's formulation had it, all men are subject to castration only because one man, the primal father, wasn't. For the masculine subject, the law is founded upon the point of its transgression. The masculine organization thus tends to be oriented towards prohibition and the fantasy of its exceptional violation. The feminine position, by contrast, asserts the totality of the universal without reference to the exception, but with the crucial proviso that the whole is internally incomplete, or non-all. Lacan relies on these two options, what might be called external exception versus internal excess, to posit two forms of enjoyment, so-called 'phallic *jouissance*' versus the '*jouissance* of the Other'. Put in crude terms, and those only narrowly relevant to sex

itself, Lacan's view here helps explain why masculine desire so typically shows a fetishistic focus on the part object, a tendency to immerse itself in the brute details of the sex act, while the feminine subject remains attuned to the larger context and situation, sensitive to a whole milieu of seduction. On the basis of this structural difference, Lacan posits an irresolvable antagonism between the sexes, tersely expressed in his dictum: *il n'y a pas de rapport sexuel*, 'there is no sexual relation'.

The relevance of Lacan's theory of sexuation to the problem of religion can be sensed in a provisional way when we remember that ancient religions conspicuously and almost universally identify the origin and order of the universe with sexual difference and sexed reproduction. The religious mind, it seems, cannot resist conflating cosmos with coitus. But the relevance becomes more interesting when sexuation is related to Christianity. Once again, we hear an echo of Kant's approach to religion in so far as Christianity turns out to be especially, perhaps even uniquely, reflective of the structures of subjectivity. The dichotomy between phallic and feminine *jouissance* as Lacan defines them can be said to centrally inhabit Christianity, not because Christianity is to be linked to one side or the other but rather precisely because of the way that it straddles the two.

The reason is that the figure of Christ is deeply ambiguous. From one point of view, Christ can be taken to stand for the ultimate exception, the perfected human subject by whose measure all others are to be judged. The Christ-exception stabilizes the law––not, to be sure, by the negative measure of transgression but rather by offering a positive exemplar of perfection. This difference, according to which the role of exception in the grounding of the law is inverted from its more familiar form of a violation or transgression of the law's moral demand, is anything but trivial. In fact, it is by means of this inversion that Christianity achieves a universalization of moral failure, establishing a perfect democracy of sin. Compared with the Messiah, we are all sinners. It is also by this means, as Nietzsche recognized, that Christianity can be thought to bring to completion the Judaic invention of bad conscience (Nietzsche 1887/1992: 493–532). Nevertheless, the logical form of the relation between the universal and the exception remains the same as that described by the masculine formula of sexuation. The universal is founded by the exception.

According to this first reading, Christianity repeats a completely traditional assertion of God as furnishing a master signifier. The organization of the Church as a hierarchy of authority that defends doctrinal orthodoxy would seem to follow with perfect consistency. And yet, a second point of view leads us to an opposite conclusion, as Christ can also be taken to instantiate the ultimate denial of all exception. The figure of Jesus inaugurates a new and unprecedented universalism that is based on his identifying himself with the human. What makes this second view distinct from the first is the way it extends the full implications of the revolutionary Christian notion of incarnation by asserting a kind of infinity inhabiting every human being, a radical equality of all persons in the eyes of God that is born of the fact that every human incarnates something of the divine. This new assertion of the universal, hinting as it does at the sort of sublime unity of the divine and the human alluded to by the mystics, is to be aligned with Lacan's feminine logic of the

non-all. Christ becomes what is 'in me more than me', a tincture of the divine. Paradoxically, a kind of essential, constitutive excess is precisely what makes me human.[20]

It is in the context of this discussion that we can make some sense of Lacan's evocation of a God who doesn't know. On the side of Christian doctrine that links Christ with the masculine logic of exception, God remains omniscient, God is fully *sujet supposé savoir*. 'God only knows', as the familiar saying has it. But according to the more radical potential of incarnation in which Christ-the-suffering-God is identified with all who suffer, Christianity is opened to the feminine logic of excess, an Other *jouissance* that is not knowable. The big Other becomes no longer identifiable with a subject supposed to know, but is, on the contrary, found to be internally incomplete, inconsistent, inhabited by an irradicable vacuity or gap. From this point of view, the Christian divinity ceases to be immutable or impassible. It is for this reason that Lacan's twentieth seminar continually flirts with the notion of an unknowing God, a suffering and kenotic being whose majesty consists precisely in a sublime form of weakness. Lacan thus probes a radical reassessment of divinity arrived at on the basis of psychoanalytic premises, provoking us with the conclusion that 'one can no longer hate God if he himself knows nothing' (1998: 98). As Slavoj Žižek has argued, this assertion of a God who does not know can be seen as a necessary correlate of identifying God with love.

> Only an imperfect, lacking being loves: we love because we do *not* know everything. On the other hand, even if we were to know everything, love would still be higher than complete knowledge. Perhaps the true achievement of Christianity is to elevate a loving (imperfect) Being to the place of God

> (Žižek 2002: 61).

The Abyssal Ground of Belief

Early in our discussion of the Lacanian Thing we noted the inadequacy of interpreting it in too Kantian a fashion. It now remains to complete the necessary distancing from Kant and draw out the consequences for our topic. Doing so means asking again about the fuller meaning of the notion of *das Ding*. What order of reality are we dealing with here?

It is immediately tempting to understand the Thing as a substantial entity of some sort. Given Lacan's framing of it in terms of the *jouissance* of the Other, for example, we might pose the Thing in more familiar Freudian terms, taking it as a kind of vortex of excessive libido, a volatile reservoir of surplus energy in the Other that threatens to wash over the subject and overwhelm it.[21] But any such interpretation would miss an absolutely key point. The Thing is in no way a substance, nor even an indeterminate quantum of energy. Nor is it an object. The Thing is the site of a fundamental blind spot, the subject's primitive acknowledgement of a zone of something unknown. The spectre of the Thing results from a kind of primitive projection which essentially says 'there is something here,

I know not what'. The Thing is thus a reality for the subject alone. It marks the space of an open question. As Freud said of it, the establishment of such a space of the question is what allows the human being to learn to cognize. As such, it performs an absolutely fundamental function, something that might be fruitfully compared to what Heidegger called the disclosive opening of human *Dasein*. It is not itself a being but is rather the lighted clearing in which beings come to presence.

The act by which this site of something unknown is posited cannot easily be named. Indeed, to call it 'posited' already implies too conscious and deliberate an exertion of the part of the subject. But even more difficult is to determine what exactly is posited. The Thing is less an object than a locus, a *topos*. *Das Ding* is not an object, nor is it a fantasy. It is rather the very frame of fantasy, the point of departure or even the cause of fantasy. The Thing is the open question to which fantasy attempts to pose an answer. As such, the Thing is the decentred void that forms the nucleus of the unconscious. One thinks of the 'other scene' that Freud attributed to the dream. In Lacanian terms, the locus of the Thing is 'extimate', something of the most intimate concern to the subject, but located outside, in the Other.

It also remains to clarify the chronology of the Thing. If it is not an object neither is it any sort of prehistorical givenness, some primal point of departure that kicks everything else into motion. Its 'being' is rather an effect of the signifier. The Thing arises with the signifier and is nothing but its excessive dimension. All of which is to say that the Thing is a kernel of the real in the Lacanian sense. And if the real and the symbolic can never be coincident, neither are they separable. They arise together. The symbolic, the real, the subject and the Other constitute an irreducibly four-fold structure. It is Lacan's answer to Heidegger's four-fold of earth and sky, mortals and divinities, and it bears the same consequence of asserting the inescapability of time.[22] This irreducible quadratic dimensionality of the four terms implies that all four are ineluctably caught up in contingent and historical specificity. Their conjunction is historicity itself.

To grasp the full implications of this precision of our discussion is to appreciate the degree to which Lacan's notion is finally less Kantian than Hegelian. Just as Hegel rejected Kant's appeal to a noumenal realm wholly beyond phenomena, Lacan insists that the real is not some impenetrable substance wholly beyond the symbolic but rather arises only from gaps, inconsistencies, and failures internal to the symbolic itself.[23] And we can now make clear how Lacan's assessment of religion parallels that of Hegel, who sharply divides the Christian legacy. On the one hand is the way that Christian incarnation, the Pauline 'scandal for the Jews and folly for the Greeks', opens up an abyssal short-circuiting of the mortal and the divine, the finite and the infinite, even to the point of evoking the death of God. On this side, Christianity marks the truly revolutionary moment in the history of human freedom. On the other is the 'positive', dogmatic aspect of Christianity, its elaboration of doctrinal and theological orthodoxies, which occasioned Hegel's critique in some of his early writings, a critique that resonates with the Pauline dictum that 'the letter killeth but the spirit giveth life' (Hegel 1795/1979).

A Lacanian perspective shows with special clarity why this division is virtually inevitable. It is rooted in the bifold function of the signifier, divided as it is between meaning (corresponding to the linkage of the signifier to a particular signified) and an indeterminate potentiality-for-meaning (the open space of a yet-to-be-determined reference, the space of the Thing). At many points, Lacan emphasizes the first theme, as when, in his interview on 'Triumph of Religion', he characterizes religion as endless spewing out of meaning. The frenzied production of meaning is both what makes religion most seductively attractive and also what makes it so inimical to psychoanalysis. 'Humanity will be cured of psychoanalysis', he says, 'by drowning the symptom in meaning, in religious meaning naturally, people will manage to repress it' (Lacan 2013: 67). But at other times, Lacan appreciates the way in which the religious impulse touches upon the real, a feature that he associates, as we have done here, with the psychic reality of the Thing. As Lacan said of it, 'Freud left us with the problem of a gap once again at the level of *das Ding*, which is that of religious men and mystics' (1992: 100).

In this way, Lacan reveals Christianity as the religion most spectacularly divided between positivity and mystery, dogma and *deinos*.[24] The Christian Church ferociously enforces doctrinal fidelity and fanatically persecutes heresy precisely because the Christian message so radically opens towards something abyssal.[25] The Christian God at once embodies the monstrosity of the Thing, the primary locus of the real, and also the symbolic big Other that seeks to defend against it. Yet the final twist consists in Lacan's refusal to allow any simple opposition between the spirit and the letter. Indeed, the keynote of his entire outlook is to insist on the primacy of the letter, along with all of the *Sturm und Drang* it brings in its train. Without the agency of the signifier, there is no spirit. As he puts it in one of his *Écrits*: 'Of course, as it is said, the letter kills while the spirit gives life. I don't disagree […] but I also ask how the spirit could live without the letter' (Lacan 2006: 423).

References

Balmés, F. (2007). *Dieu, le sexe et la verité*. Paris: Erès.

De Beauvoir, S. (1989). *The Second Sex*. New York: Vintage Books.

Boothby, R. (2001). *Freud As Philosopher: Metapsychology After Lacan*. New York: Routledge.

Chiesa, L. (2016). *The Not-Two: Logic and God in Lacan*. Cambridge, MA/London: MIT Press.

Copjec, J. (2015). 'Sex and the Euthanasia of Reason'. In *Read My Desire: Lacan Against the Historicists* (pp. 201–236). London/Brooklyn, NY: Verso.

Daviot, P. (2006). *Jacques Lacan et le sentiment religieux*. Paris: Erès.

Depoortere, F. (2007). 'The End of God's Transcendence? On Incarnation in the work of Slavoj Žižek'. *Modern Theology 23*(4) October: 497–523.

Freud, S. (1886–99). *Pre-Psycho-Analytic Publications and Unpublished Drafts*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 1. London: Hogarth Press.

Freud, S. (1909). *Two Case Studies*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 10. London: Hogarth Press.

Hegel, G. W. F. (1795/1979). *Early Theological Writings*. Philadelphia, PA: University of Pennsylvania Press.

Heidegger, M. (1962). *Being and Time*. New York: Harper and Row.

Lacan, J. (1968). 'Seminar XVI: From an Other to the other', translated for private use from the unpublished French transcript of the seminar by Cormac Gallagher.

Lacan, J. (1981). *The Four Fundamental Concepts of Psycho-analysis*. New York: W. W. Norton.

Lacan, J. (1992). *The Seminar of Jacques Lacan, Book VII, The Ethics of Psychoanalysis*. New York: W. W. Norton.

Lacan, J. (1998). *The Seminar of Jacques Lacan, Book XX, Encore: On Feminine Sexuality, The Limits of Love and Knowledge*. New York: W. W. Norton.

Lacan, J. (2006). *Écrits*. New York/London: W. W. Norton.

Lacan, J. (2013). *The Triumph of Religion*. Cambridge/Malden, MA: Polity Press.

Lacan, J. (2014). *The Seminar of Jacques Lacan, Book X: Anxiety*. Cambridg/Malden, MA: Polity Press.

Lacan, J. (2017). *The Seminar of Jacques Lacan, Book VIII, Transference*. Cambridge/Malden, MA: Polity Press.

Leupin, A. (2004). *Lacan Today: Psychoanalysis, Science, Religion*. New York: Other Press.

Nietzsche, F. (1887/1992). 'The Genealogy of Morals, Second Essay "Bad Conscience"'. In *Basic Writings of Nietzsche* (pp. 493–532). New York: Modern Library.

Pound, M. (2007). *Theology, Psychoanalysis, and Trauma*. London: SCM Press.

Richardson, W. (1997). '"Like Straw": Religion and Psychoanalysis'. *The Letter: Lacanian Perspectives on Psychoanalysis 11*: 1–15.

Žižek, S. (2000). *The Fragile Absolute—or, Why is the Christian Legacy Worth Fighting For?* London/New York: Verso.

Žižek, S. (2001). *On Belief*. London/New York: Routledge.

Žižek, S. (2002). 'The Real of Sexual Difference'. In S. Barnard and B. Fink (eds), *Reading Seminar XXL Lacan's Major Work on Love, Knowledge, and Feminine Sexuality* (pp. 57–75). Albany, NY: State University of New York Press.

Žižek, S. (2003). *The Puppet and the Dwarf: The Perverse Core of Christianity*. Cambridge, MA/London: MIT Press.

Žižek, S. (2005). 'Neighbors and Other Monsters: A Plea for Ethical Violence'. In S. Žižek, E. Santner, and K. Reinhard, *The Neighbor: Three Inquiries in Political Theology* (pp. 134–190). Chicago, IL: University of Chicago Press.

Žižek, S. (2006). *How to Read Lacan*. New York: W. W. Norton.

Žižek, S. and Gunjevic, B. (2012). *God in Pain: Inversions of Apocalypse*. New York: Seven Stories Press.

Žižek, S. and Milbank, J. (2009). *The Monstrosity of Christ: Paradox or Dialectic?* Cambridge, MA/London: MIT Press.

Notes:

(1.) The past decade has seen a flurry of commentary on the implications of Lacanian theory for rethinking religion. Among the more notable are Pierre Daviot (2006), *Jacques Lacan et le sentiment religieux*; François Balmes (2007), *Dieu, le sexe et la verité*; Marcus Pound (2007), *Theology, Psychoanalysis, and Trauma*; and Lorenzo Chiesa (2016), *The Not-Two: Logic and God in Lacan*. But no one has done more to demonstrate how Lacan reopens the field of the religious from a psychoanalytic point of view than Slavoj Žižek. Žižek treats questions of religion and God in many of his books, perhaps most directly in *The Fragile Absolute—or, Why is the Christian Legacy Worth Fighting For?* (2000), *On Belief* (2001), *The Puppet and the Dwarf: The Perverse Core of Christianity* (2003), *The Monstrosity of Christ: Paradox or Dialectic?* (with John Milbank 2009), and *God in Pain: Inversions of Apocalypse* (with Boris Gunjevic 2012).

(2.) [Eds. See Cottingham, this volume, and Blass, this volume, for more detailed discussions.]

(3.) [Eds. See Eagle, this volume, §'Falsifiability of psychoanalytic propositions', for a discussion of interpretations of the Oedipus complex in relation to empirical testing.]

(4.) 'God is real' states the title of the final chapter of Leupin (2004). For other treatments of God as an embodiment of the real, see Depoortere (2007: 497–523) and Richardson (1997: 1–15).

(5.) For a clear, freewheeling, and highly suggestive introduction to the Lacanian concept of the big Other, see Žižek (2006).

(6.) [Eds. For a very different account of the need for meaningfulness in relation to God, see Cottingham, this volume, §'Freud's critique of religion'.]

(7.) My thesis, which cannot be developed adequately in this chapter, is that while explicit references to the Thing tend to disappear after the seventh seminar, the underlying notion is not abandoned but rather replaced by its representative, the so-called *objet petit a*. When, late in his career, Lacan referred to the *objet a* as his greatest single contribution to psychoanalysis, we should also discern an implicit reference to the concept of the Thing.

(8.) I capitalize 'Other' here and will continue to do so throughout this chapter, but the choice is an awkward one in so far as 'Other' must do double duty between the concepts of the little and big Others. In fact, Lacan himself alternates in his capitalization of *Autre* throughout his work without any apparent consistency. The most logical thing would seem to be using the lower case for the little Other and the upper case for the big Other. But then again, even the little Other of the fellow human being sometimes deserves the emphasis lent by the capitalization, precisely because, when its Thingly dimension is taken into account, the fellow human being becomes something totally different than the impression of ordinary experience leads us to conclude. It is to recognize this point that I will retain the capitalization even of the 'little Other'.

(9.) Linking the primal root of anxiety with the mother herself makes new sense of Freud's observation that the German word '*heimlich*', which by virtue of its reference to 'home' (*Heim*) indicates what is safest and most familiar, but may also, according to its secondary meaning as what remains 'concealed, kept from sight, so that others do not get to know of or about it' function as an equivalent for what is '*unheimlich*, or 'uncanny'. Cf. Freud, (1955: 224).

(10.) 'In the body there is always, by virtue of this engagement in the signifying dialectic, something that is separated off, something that is sacrificed, something inert, and this something is the pound of flesh.' (Lacan 2014: 219).

(11.) 'The most decisive moment in the anxiety at issue, the anxiety of weaning, is not so much when the breast falls short of the subject's need, it's rather that the infant yields the breast to which he is appended as a portion of himself.' (Lacan 2014: 313).

(12.) 'Here the *a* stands in for the subject––it's a stand-in in the position of precedent. The primordial, mythical subject, posited at the outset as having to be constituted in the signifying confrontation, can never be grasped by us, and for good reason, because the *a* preceded it and it has to re-emerge secondarily, beyond its vanishing, marked by this initial substitution. The function of the yieldable object as a piece that can be primordially separated off conveys something of the body's identity, antecedent to the body itself with respect to the constitution of the subject' (Lacan 2014: 314).

(13.) The quotation 'hell is other people' famously belongs to Jean-Paul Sartre. By citing it here I don't mean to conflate Lacan's outlook with that of Sartre. The two are importantly different in ways we cannot expound upon here. Nevertheless, there is a significant, if only partial, overlap with Sartre when it comes to Lacan's treatment of the neighbour-Thing of the fellow human being.

(14.) The precise phrasing of 'the word as the murder of the thing', indeed most of Lacan's acquaintance with Hegel, was made known to him via the highly influential lectures of Alexandre Kojève.

(15.) [Eds. See also Macdonald, this volume.]

(16.) Cf. Lacan's remark: 'That is clearly the essence of law––to divide up, distribute, or reattribute everything that counts as *jouissance*' (1998: 3).

(17.) I have elsewhere put forward an account of sacrifice from a Lacanian point of view. See Boothby (2001: 175–89).

(18.) [Eds. For alternative accounts of the roots of religion in subjectivity, see Cottingham, this volume, and Blass, this volume.]

(19.) Lacan elaborates the formulae of sexuation most extensively in *Seminar XX: Encore*. For especially useful explications of his theory on this point, see Copjec (2015: 201–36), and also Žižek (2002: 57–75).

(20.) Cf. the final chapter of Lacan (1981): 'In You More Than You'.

(21.) We might note in passing that posing the Thing as a mass of energy makes it comparable to the Freudian id, and also invites a link with Lacan's 1955 essay on 'The Freudian Thing' (2006: 334–63) in which the id, *le ça*, is said to speak: *ça parle*.

(22.) Cf. Lacan's remark: 'I have already asked the question here as to what the critical conceivable minimum is for a signifying scale, if the register of the signifier is to begin to organize itself. There cannot be a two without a three, and that, I think, must certainly include a four, the quadripartite, the *Geviert*, to which Heidegger refers somewhere' (2002: 65–6).

(23.) Linking Lacan with Hegel is a tricky business iinsofar as Lacan himself retained a persistently critical view of Hegel as a philosopher of metaphysical closure. Here again, credit needs to be given to Slavoj Žižek, whose clarification of Hegel's thought, convincingly offering an almost complete inversion of the dominant interpretation, has made it possible to recognize the deep conjunction between Hegel's philosophy and Lacan's retheorization of psychoanalysis.

(24.) One wonders whether it is not for reason of Lacan's recognition of the particularly intensity of this double function in the case of Christianity that he claims that 'There is *one* true religion and that is the Christian religion' (Lacan 2013: 66).

(25.) This abyssal dimension is what Žižek and Milbank (2009) has called 'the monstrosity of Christ'.

Richard Boothby

Department of Phlosophy, Loyola University Maryland

Oxford Handbooks Online

Abstract and Keywords

This Handbook examines the contributions of philosophy to psychoanalysis and vice versa. It explores the most central concept of psychoanalysis—the unconscious—in relation to its defences, transference, conflict, free association, wish fulfilment, and symbolism. It also considers psychoanalysis in relation to its philosophical prehistory, the recognition and misrecognition afforded it within twentieth-century philosophy, its scientific strengths and weaknesses, its applications in aesthetics and politics, and its value and limitations with respect to ethics, religion, and social life. The book explains how psychoanalysis draws our attention to the reality of central aspects of the inner life and how philosophy assists psychoanalysis in knowing itself. This introduction elaborates on the phrase 'know thyself', the words inscribed at the Temple of Delphi, and illustrates the connection between matters philosophical and psychoanalytic in relation to the Delphic command by highlighting their mutual concern with truth and truthfulness.

Keywords: philosophy, psychoanalysis, unconscious, ethics, inner life, knowing oneself, truth, truthfulness

A Delphic Command

The words inscribed at the Temple of Delphi—'know thyself'—have been a guiding light for both philosophy and psychoanalysis.

Thus it is often said that an important aim of psychoanalysis is self-knowledge.[1] An explicit connection to Delphi, however, was made by Freud himself just once, in *The Psychopathology of Everyday Life*, where he says 'The road whose goal it is to observe the precept γνῶθι σεαυτόν [*gnothi seauton*: know thyself] runs *via* the study of one's own apparently accidental actions and omissions' (Freud 1901: 210). The claim that this was not only an aim of psychoanalysis, but the guiding principle of Freud's life and work, was left to Ernest Jones to make in his obituary of Freud (in the kind of hagiographic language which could only be justified, if at all, in such a context): 'Future generations of psychologists will assuredly wish to know what manner of man it was who, after two thousand years of vain endeavour had gone by, succeeded in fulfilling the Delphic injunction: know thyself' (Jones 1940: 4). Jones rehearsed the claim again in his influential biography of Freud (Jones 1955: 470), and the connection between psychoanalysis and the inscription at Delphi became widespread, endorsed by such luminaries as Sterba (1969: 439), Eissler (1963: 461), Menninger (1956: 623), Kohut (1973: 25), and Bettelheim (1982: ch. 4), who elaborates the aim by connecting it with practical effects: 'The guiding principle of psychoanalysis is that knowing oneself requires knowing also one's unconscious and dealing with it, so that its unrecognized pressures will not lead one to act in a way detrimental to oneself and others' (1982: 24).

While the Delphic command is mentioned in several of Plato's dialogues, Socrates, like Freud, takes it up in discussion just once, in the *Phaedrus*, where he claims that it is 'ridiculous' to pursue knowledge of other things before one knows oneself, a claim which places self-knowledge at the heart of the human epistemological endeavour. What Socrates himself meant by inviting us to pursue self-knowledge is something about which scholars disagree (Hadot 1995: 90), yet there are at least four senses in which the endeavour has been considered central to philosophy.

First we have the practical endeavour of the individual philosopher, through her philosophical reflection on her predicaments in their particularity, better to come to know *herself*. This goal, clearly apparent in the approach of Socrates, the Stoics, and Epicureans, has been suggested as the intelligibility-conferring setting against which various ancient philosophical texts must be read (Hadot 2002). An explicit personal, therapeutic, existential engagement in doing philosophy is only occasionally to be found in more recent philosophers (e.g. Nehamas 1998), yet may play an implicit motivating role for philosophers more widely.

Next we have that project which could be called 'knowing *ourselves*'—namely the philosophical project of clearly articulating the human condition in its generality. This is the Socratic task taken up in theoretical mode. This project of philosophical anthropology

may appear merely descriptive yet, to the extent that humanity is to be understood by reference to ideals, aspirations, and excellences (such as truth and truthfulness, love and goodness, reason and rationality), it is not intelligibly separable from fundamental evaluative questions of how to live (i.e. from what we here call 'ethics', which we understand more broadly than questions of 'morality' traditionally conceived).

Thirdly we meet with the philosophical valuation of knowing ourselves as an intrinsically valuable way of life. This was most famously encapsulated in Socrates' famous retort when invited at his trial to abandon reflection on the Delphic oracle's pronouncements (i.e. to abandon his search for wisdom): 'the unexamined life is not worth living'. Here philosophical reflection becomes not merely a reflective means to a non-reflective end but an end in itself, one which constitutes a significant form of the good life.

Finally we may bring these together, as when my philosophical reflection on how *I* am to live is informed by my best understanding of how *one* may live best *given what it is to be human*, or when my understanding of how *one* may live best is informed by deep reflection on *my* own (or others') personal experience. While different philosophers have understood the connection between self-knowledge and the ethical life in various ways, it is notable that living an examined life, one enabled by philosophical enquiry, has been seen as the crowning purpose and achievement of philosophy by many, from Plato and Aristotle, through Descartes, Spinoza, Nietzsche, Husserl, Heidegger, and Dewey to Martha Nussbaum and Bernard Williams. And here we meet with values that find expression not only in philosophy but also in the goals and methods of psychoanalysis. In what follows we explore the character and congruence of such endeavours at self-knowledge.

As Bernard Williams expresses our second, theoretical, version of the Delphic command, philosophy can be understood as 'part of a more general attempt to make the best sense of our life, and so of our intellectual activities, in the situation in which we find ourselves' (2000: 479). As such, he argues, philosophy is a 'humanistic discipline'. It is humanistic not only in the sense that the central object of study is ourselves, but also that its understanding develops within and expresses a 'human perspective'. Definitive of such a perspective is that it is irreducible to that of the natural sciences in its style, method, and aims. The scientific ambition that particularly concerns Williams is that of working towards a description of the world 'as it is in itself, independent of perspective' (2000: 481). The contrastive aim of philosophy, as he understands it, is to do proper justice to matters of meaning, intelligibility, and significance—matters which, he argues, involve reason reflecting on itself from within, drawing inescapably on such perspectival modes of understanding as are inevitably historically and culturally situated and conditioned.

Williams's thought may here helpfully be brought into relation with that of Charles Taylor (1985) who takes us closer to matters psychoanalytic with his focus on that project of self-understanding which involves our articulating—both in the sense of 'giving voice to' and in the sense of 'developing and refining'—our *emotional* experience. Thus according to Taylor we are essentially 'self-interpreting animals' whose 'interpretation of ourselves

and our experience is constitutive of what we are, and therefore cannot be considered as merely a view on reality, separable from reality, nor as an epiphenomenon, which can be by-passed' (1985: 47). What Taylor means here by 'self-interpretation' is not an essentially intellectual or reflective act, but nevertheless pertains to a kind of self-understanding not met with in animal life. We highlight two aspects of Taylor's discussion.

First, Taylor argues that 'experiencing a given emotion involves experiencing our situation as bearing a certain import' (1985: 49). An import is a way in which a situation or object can be relevant and important to us, given our desires and purposes. Taylor's claim is now very widely accepted in the philosophy and psychology of emotion as the claim that emotions constitutively involve our making appraisals of situations, which appraisals form 'the grounds or basis for the feeling' (49). Thus, we can define emotions 'by the imports they relate to: fear is the affective response to the menacing, anger to the provoking, indignation to the flagrantly wrongful, and so on' (49). As a consequence, our emotions are not intelligible for what they are merely in objective—i.e. experience-independent, physiological, or physical-causal—terms since they essentially 'characterize things in their relevance to our desires and purposes' (51).

Second, Taylor highlights the relationship between our self-understanding, our social and moral emotions, and what it is to lead a human life. Shame, for example, is only intelligibly experienced by those who understand, experience, and value their lives as having a certain kind of import—i.e. only intelligibly experienced by subjects with an aspiration to dignity (1985: 53). The subject who feels shame cares about how she handles herself and how she is seen by others. She is someone who understands herself, in her shame, to be failing to meet standards that matter to her. In such ways are self-conscious social subjects partly constituted, in their emotional lives, by their self-interpretations.

None of this is to say that there is no space for our self-interpretations to go awry, nor to suggest that we may not fail to interpret ourselves in apt ways (see Moran 2001: ch. 2; Carman 2003). What it suggests, however, is that: from the point of view of philosophical anthropology, aspirations to grasp what it is to lead a human life in merely objective terms (i.e. in terms not referring to a subject's self-understanding) will be doomed to failure; from the point of view of psychopathology, that the relationship between self-interpretation and selfhood must be taken into account in understanding the distinctive sufferings of human selves; and from the point of view of therapy, that we can begin to understand how a merely talking cure could be thought curative of such disturbances as reach down into our selfhood—since reinterpreting the meaning of one's behaviour will be at the same time a refashioning of oneself.

Recovering the Inner Life

Williams and Taylor pursue philosophy as, in part, a project of retrieving, sustaining, and exploring a conception of what it is to live a human life. The living of a distinctly human life involves, at its paradigmatic best, the experience of a rich inner subjectivity which informs our relationships, work, and creative projects. Aspects of this inner life may, and frequently do, remain stunted or atrophy through lack of nurture, become covered over by emotional defences against the pains of living, be inhibited by ideological narrowness, or be degraded through corrupt forms of relationship. Against such natural tendencies, many of the arts and humanities, including literature, poetry, history, and philosophy, engage in a continuing project of developing and retrieving our subjectivity. The project may be valuable for individuals yet also, more broadly, for cultures.

What makes this ongoing endeavour intelligible to its participants is the possible disjunction of our *implicit, lived* understanding of what it is to live a distinctly human life and those more *explicit* self-understandings which may do better or worse justice to what is implicitly grasped and lived. What makes it necessary is the way in which our implicit, lived understandings may have been muted or thwarted by various factors including our explicit self-conception that may tacitly or explicitly trash much of what makes for our humanity. (Imagine how impoverishing it would be to actually live as if that philosophy or psychology of mind you consider most implausibly reductive, or which simply leaves out of account or sees as merely epiphenomenal that which in your inner and existential life most matters to you, were true.) And what makes it valuable is the recovery of that within us which has become muted, our emancipation from such falsifying self-understandings as impoverish our self-becoming, and the intrinsic dignity of the examined life.

Nothing in such a humanistic approach is opposed to the natural scientific study of human life, but it challenges the scientistic thought that such a life may itself be adequately articulated in the terms offered by the natural sciences. The humanity that Williams and Taylor retrieve for us is one run through with a constitutive normativity, subjectivity, affectivity, rationality, and agency, and the methods they retrieve for us are ineliminably hermeneutic and aspire to no interpretation-transcending objectivity. What matters for them is not that the humanities and social sciences demonstrate the validity and reliability of their methods by adopting natural scientific methods ill-fitted to self-interpreting animals, but rather that they self-critically deploy such meaning-apprehending methods as are apt to the study of human life.

Such a focus is also central to the vision of psychoanalysis which pursues its own exploration and development of our humanity on two fronts.

First, at the level of theory, and by contrast with behaviourist, cognitivist, and physiological psychology, it emphasizes the inner subjective life. This is the life of our preoccupations—with erotic desire and social recognition, with our shame, guilt, and contrition, power and humiliation, with our hopes and our histories, lovableness and

loneliness, lovesickness and consuming hatreds, shyness and courage, envy, resentment, and gratitude, intense secret passions, idealizing delights, the peculiarly vaunted or denigrated status for us of our significant attachment figures, our sexual adequacy and inadequacy, fateful repetitions, with our expressions and deeds which threaten to betray us, and all of our inner conflicts, moodiness, anxieties, excitements, self-punishments, self-defeating behaviours, irrational impulses, and bodge-job forms of self-management. This caboodle is what we may call our 'subjective life' or 'internal world', and many a student of psychology has been disappointed to find that what they naturally hoped would be centre stage on their syllabus—namely why our emotional life is so often baffling and tumultuous—barely gets a look-in besides the studies of cognition, behaviour, perception, and neuropsychology. It is, of course, the life of the neurotic subject—but also of *all* of us, since 'psychoanalytic research finds no fundamental, but only quantitative, distinctions between normal and neurotic life' and 'the psychical mechanism employed by neuroses is not created by the impact of a pathological disturbance upon the mind but is already present in the normal structure of the mental apparatus' (Freud 1900: 373, 607).

Second, at the level of practice, psychoanalysis aims to attain or recover for us the very subjective sense of what otherwise appears not as meaningful, humanly intelligible moments of emotionally charged behaviour, but instead merely as behavioural signs and symptoms of an unknown condition. It aims, that is, to restore or develop in us our subjectivity, to help us recover or grow our agency or self-possession, to retrieve or awaken our inner lives, to 'make the unconscious conscious' and thereby to 'know ourselves'.

It is safe to say that, compared to any other psychological school, psychoanalysis has in both its theory and practice most keenly kept its pulse on the distinctive qualities of the inner life. That philosophy may borrow from it to considerably enrich its own sense of what it really means to live a distinctly human life should not be surprising (e.g. Wollheim 1984; 1993). One of the most valuable contributions made by psychoanalysis to the project of making sense of ourselves is its drawing our attention to both the clinical data and the everyday observations upon which it constructs its theories. It offers up, if not an entirely new, then a considerably under-examined, set of human experiences. Such experiences are relevant not only for psychoanalysis' own explanatory and therapeutic projects, they also deserve a place in the understanding of what it is to be human at play in many other disciplines—including, of course, philosophy. Many chapters in this Handbook consider the significance of psychoanalysis as a contribution to the meaning and meaningfulness of human activity, to the nature of human experience, to a philosophical anthropology and the phenomenology of human consciousness and relating, and thus to questions in ethics, religion, aesthetics, and, of course, self-knowledge.

Thus, psychoanalysis draws our attention to the reality of central aspects of the inner life which we know implicitly to be essential to human life as lived yet which for various reasons often escape our reflective grasp. As Nietzsche remarked, philosophy is littered with claims and ideas, e.g. about human psychology and ethics, that are insufficiently tied to human reality. For example, the important Aristotelian conception of man as a 'rational

animal' might, if we're not careful, illegitimately displace from our self-understanding the essential contribution made by our emotional sensibilities, sensibilities which make possible not only irrationality and human impoverishment but which put us in contact with reality and enable our flourishing. Or an equation of mentality with consciousness may squeeze out of view the essential contribution to our psychological lives of dynamically and descriptively unconscious mental processes. At the time of writing, many areas of philosophy are undergoing transformation in response to developments in the social, cognitive, and neuropsychological understanding of unconscious processes and the possible challenges these provide to the autonomy and integrity of conscious rational deliberation. The issue is at the heart of philosophy's project, as the place and nature of reason and conscious deliberation have been of central concern to philosophy since its inception.

A second reason why philosophy should attend to the understandings of human life offered by psychoanalysis, and may be enhanced by psychoanalytic reflection, is that the unconscious may be understood to consist of optional and idiosyncratic aspects of our lives which go unexamined and constrain our sense of what is possible (see Fuchs, this volume, and Lear, this volume). This point can be better understood in light of a more familiar argument concerning why philosophy should attend to history, deriving from the self-interpreting nature of human life. If what it is to live a human life or have a human mind were immutable facts, they could be interrogated by means of a familiarity with any human culture at any point in history (one's own culture and times, for example). Since, however, what it is to be a human being itself changes (within limits) over time and place, philosophers attempting to grasp what it is to live a human life or to be minded in human ways will do well to attend to more than their present time. This will be important not only for the understanding of other modes of human life but, perhaps even more importantly, for the understanding of our own. For it is only when set against the backdrop of other ways of being human that we can understand the distinctive shape, and acknowledge the contingency, of our own life. This lesson from history, we say, has a psychoanalytic analogue given above. Thus, a historical, sociological, *and* psychoanalytic method may help philosophy come to know itself, to make its unconscious conscious, by unearthing the contingent character of the forms of life which it takes for granted, including the form of its own enquiries.

Psychoanalysis: Know Thyself

But in what ways may philosophy return the favour? Various chapters of this Handbook provide their own diverse answers. But, following here the theme of 'know thyself', we propose that one of philosophy's distinctive contributions is to assist psychoanalysis in knowing *itself*. Psychoanalysis as a discipline involves an ongoing dialectic of psychological theory and therapeutic practice informing each other, over time generating diverse self-understandings that are in tension with one another. Perhaps the debate over

natural science and hermeneutics provides the most striking example of this. Philosophy can make more explicit conceptual implications, uncover misunderstandings, unearth problematic structuring assumptions, and enable new and productive self-understandings.

As an illustration: it is today fairly well understood that the kind of self-knowledge sought both by psychoanalysis and by anyone hoping to do genuine emotional work on herself, is not simply reflective. It may be interesting for us to develop beliefs about our minds' functioning, and such beliefs may even be true, but, from a therapeutic point of view, the risk is significant that such an intellectual self-acquaintance may defensively stand in the way of, and disguise the ongoing need for, an emotionally deeper and mutative form of self-knowledge. Relatedly, Jonathan Lear (2005: section 4) notes that even 'raising "the question of how to live" can be a way of avoiding the question of how to live'. Getting clear on what is and isn't involved in that deeper and intrinsically transformative project of knowing thyself is something with which philosophy has been concerned for centuries before psychoanalysis, and without philosophical aid it is inevitable that psychoanalysis will sometimes embed those self-misunderstandings about what it is to know thyself with which philosophy has had to grapple.[2]

Consider for example Hadot's (1995: 90) explication of that first sense of 'know thyself' offered in Section 1:

> [In] the Socratic dialogue … the interlocutors are invited to participate in such inner spiritual exercises as examination of conscience and attention to oneself; in other words, they are urged to comply with the famous dictum, 'Know thyself'. Although it is difficult to be sure of the original meaning of this formula, this much is clear: it invites us to establish a relationship of the self to the self, which constitutes the foundation of every spiritual exercise.

But is 'this much' clear? Sometimes this is the relationship taken in knowing oneself, but others have taken the Delphic command otherwise. Augustine, for example, (*c*.417/2002: 10.8.11) urged that when the mind 'is ordered to know itself, let it not seek itself as though withdrawn from itself, but let it withdraw what it has added to itself'. That is to say that some central forms of knowing oneself involve not the gleaning of new information about oneself—not an addition but a subtraction, not so much the establishment of a helpful self-relation but the undoing of an unhelpful self-relation. In addition we also meet with those other forms which take us closer to self-becoming, i.e. to growing into one's own character, to articulating and making more determinate what is as yet undeveloped, than to any increased store of knowledge about oneself.

Philosophical reflection on the way that the term 'self' works in a variety of locutions—including 'know thyself'—bears this out and helps us guard against assuming that it *inevitably* signifies the object of a reflexive relationship, and instead helps us see how it functions to signal the absence of relationship. Consider, for example, 'self-respect', 'self-possession', 'self-motivation', 'self-consciousness', 'to thine own self be true', 'selfishness'. Someone who 'selfishly' 'keeps something for himself' is not well

understood as keeping something for someone who happens to be himself, but rather as simply keeping something *without thought of later giving it to* anyone. The same may be said of someone who 'keeps something *to* himself': he just *doesn't share it with* others. Someone who is 'self-possessed' doesn't stand in a positive relation of possession to himself (whatever that would mean), but is rather *free from the psychological influence of* others past and present. He now acts in a straightforward and decisive manner that is free from responsibility-avoiding dither. Someone who becomes 'self-conscious' is, to be sure, thrown into an anxious state of wondering how *she* is coming across to others, and in this sense it is she and they, these flesh and blood people (but not some additional 'self' that she has), who are the objects of her attention. Yet what is also essential to our self-conscious subject is that she has been *thrown out of relationship* with these others, at least in any trusting connected form. Someone who is 'true to herself' is not so much simply representing herself accurately or simply happening to act on the basis of whatever she desires. Rather, she is *not in the business of dissimulation* but now *chooses and acts* according to her own values. The 'self-motivated' person is simply a person whose motivation to achieve her goals is *not dependent on external influences*. And, we suggest, one important understanding of the person who follows the Delphic command focuses not on the subject enjoying a reflexive relationship with his own mind but on his *enjoying the absence of unhelpful defence mechanisms*. For just as 'being true to oneself' is not perspicuously taken to cover, say, unremarkable cases of wanting to go for a walk and then, by gosh, going for a walk, so too 'knowing thyself' is not perspicuously taken to cover, say, ordinary cases of being able to verbally express one's thoughts and feelings about, say, going for that walk. Instead 'knowing thyself' is, in such cases, not a matter of having, but rather of *emancipating oneself from*, a certain relation with oneself—a relation of self-deception or the inability to tolerate and own one's thoughts and feelings. In such contexts, at least, the injunction to 'know thyself' refers not to the cultivation of a truthful reflexive presentation of the mind to itself but to a non-dissimulative, non-reflexive engagement with one's life.[3]

Or—by way of understanding the value of philosophical reflection on 'knowing thyself' to psychoanalysis—consider the idea that such self-knowledge is by itself but a morally neutral affair. Here we are offered the notion of a psychopathic patient who, despite her depravity, and perhaps even because of a successful psychoanalysis, knows herself perfectly well—not just in the sense of having a factually correct understanding of her own psychology, but also in the sense of having dismantled her defences. Without philosophical reflection such an idea can seem rather too clearly intelligible, as if there were no meaningful alternative. Yet, while the claim remains controversial, it certainly has been doubted whether what we understand by the Delphic command, or for that matter by the work of psychoanalysis, can quite so readily meet with an amoral interpretation. For, the suggestion goes, just as it is only in and through our joy, anger, and sadness that we really understand what it is to know reward, be wronged, and lose what matters to us, so too it is only really in and through relationships conditioned by love that we can comprehend others' true reality, and only in and through our honest commitment, integrity, and humility that we can know what it is to lead a truly human

life. On this view we do not grasp what it is to be a human living a distinctly human life by becoming knowledgeable about the behavioural habits of *Homo sapiens*. Rather we grasp it by engaging in forms of relationship with others which are essentially characterized in moral terms—in terms of what is humane, in terms of what is revealing of our and their humanity.[4] Truly understanding what it is to wrong someone, for example, may be thought to consist not simply in being able to track a range of abstract propositional entailments, but instead in feeling guilt and wanting to put it right. There are certainly senses in which a clever and uninhibited psychopath may 'know herself', but if she can't feel the guilt she has nevertheless accrued by her evil acts then, the suggestion goes, there is yet an important form of self-knowledge which she lacks.

This excursus into two important, psychoanalytically pertinent meanings of 'knowing thyself' is of course but one of the ways in which philosophy may repay psychoanalysis for its enriched reflective conception of our inner life. Stepping back to survey the interdisciplinary field we may distinguish advocative, critical, and synoptic applications of philosophy to psychoanalytic theory.

On the advocative side we find philosophy helping to defend and clarify psychoanalytic theory from misunderstandings. Here we might think, for example, of how best to understand the physicalistic and energetic metaphors within psychoanalysis, how best to understand the notions of psychological 'structures' and defence 'mechanisms', and which methods of investigation, epistemic standards, and forms of understanding and explanation are best suited to knowledge of inner life.

On the critical side we find philosophy helping to sort out the wheat from the chaff within psychoanalytic theory. What has mattered here is not so much the scientific evidence for the truth or falsity of psychoanalytic theories, which is not a direct matter for philosophy. Instead what matters here is the cogency of the forms of reasoning used within psychoanalysis to support its claims, and the critical unearthing of optional, perhaps even worrisome, moral and political values tacitly embedded in the theory and practice. Here, connecting the critical with the advocative application, the question of whether psychoanalysis can qualify as a science, and if so, what kind of science, has been of considerable importance. In such ways psychoanalysis comes to better know itself—to know and work through its habitual irrationalities in the service of achieving a more honest, perhaps a more modest, less hubristic, and more integrated enterprise.

The synoptic contribution of philosophical reason to psychoanalysis takes us into a broader discussion of interdisciplinarity in psychoanalytic thought.

Psychoanalysis Situated

The earlier discussion of the Delphic command stressed the importance of not misunderstanding 'knowing thyself' as always and inevitably involving the cultivation of a positively informative relation of the self to itself. Instead it urged the significance of a form of 'knowing' marked principally, not by the presence of a particular epistemic attitude but rather, by the *absence* of self-deception. Another way to misunderstand 'knowing thyself' merely as a self-relation would be to assume that it did not essentially implicate others—i.e. to overlook the essentially *interpersonal* nature of self-knowledge. This idea that we come to know ourselves in and through one another forms the heart of Hegel's (1807/1977) conception of identity as 'negation': we become who we are in so far as we distinguish ourselves from others and in so far as we achieve mutual recognition with them.[5] Such negation determines not only our sensorimotor selfhood—as we differentiate from, while at the same moment we perceptually and motorically relate to, our proximate environments—but also our personalities—as we come to relinquish our childhood egocentricity. In that process we come to appreciate both that we and others have genuinely different tastes, desires, and values and that, all being well, for all that we may still offer one another humane recognition. Or, at least, that we may—if we so wish— work to achieve that mutual recognition and mutual accountability, work to manage our relationships, and overcome our and others' misrecognitions, such unending work being —in the 'tragic vision' of psychoanalysis—an aspect of any worthwhile relating at all.

That self-knowledge may be won through essentially relational means marks a significant theme of recent psychoanalysis which has come to place our object relations and our intersubjectivity at the heart of both its clinical theory and its clinical practice. Yet negation and relation also provide another means for philosophy to repay its debt for the richer picture of the inner world offered to it by psychoanalysis: by helping psychoanalysis 'know itself' through grasping its relations to, identities with, distinctions from, debts to, and dependencies on other disciplines. Here philosophy plays its long-established role of coordinating and synthesizing synoptician.

Psychoanalysis itself originated in interdisciplinary reflection. It is both well known and frequently forgotten within psychoanalysis that Freud drew upon extensive non-clinical sources in constructing psychoanalytic theory. Patricia Kitcher (1992) lays out the full extent of Freud's borrowings from theories and discoveries of his time in neurophysiology (neurons, psychic energy, the reflex model of the mind), psychology (associationism, functional analysis), psychiatry (unconscious ideas, the sexual origin of neurosis, the separation of ideas and language), sexology (infantile sexuality, stages of sexual development, component instincts), anthropology and evolutionary biology (recapitulationism), with further ideas taken from philology and sociology. To the extent to which psychoanalysis draws on non-clinical findings and models directly it will need to revisit such ideas as have been superseded within their source disciplines. The same may be said of the use by more recent incarnations of psychoanalysis of theories and concepts

from structural linguistics, attachment theory, existential phenomenology, Marxism and critical theory, anthropology, postmodernism, developmental psychology, and neurobiology. Philosophy in synoptic, grand-theoretic, mode may take up the task of urging and facilitating such updatings, and of drawing critical attention to failures to do so.

Philosophy's job here, however, is not only the uncomfortable one of interdisciplinary police officer, but also that of diplomat. For sometimes, when psychoanalysis borrows from other disciplines, it does not so much directly import their concepts, as tacitly reappropriate or metaphorically extrapolate them for its own ends. While this may, in many cases, relieve psychoanalysis of the obligation to keep track of changes in scientific knowledge and understanding within the source domains, it may result in unclarity about the imported concepts, e.g. whether they are best understood as carrying literal or metaphorical senses (an example here may be Freud's use of energetic and biological concepts). Here the task of philosophy is to clarify this indeterminacy and to assess whether inferences are being made within the psychoanalytic theory which illegitimately switch the senses of terms mid-argument.

Perhaps the most significant diplomatic role for philosophy concerns the questions of whether and how the findings and the methods of non-psychoanalytic disciplines are to be brought to bear on psychoanalytic theory and vice versa. Looking back a few decades one thinks especially of attempts to use the quantitative methods of experimental psychology to test hypotheses derived from psychoanalytic theory, or to test the adequacy of psychoanalytic therapy. The distinctly philosophical questions here were whether and when and how such hypotheses are pertinent to the theory, whether the theory is genuinely testable, whether it's too bad for the theory or too bad for the experimental methods if it isn't, and whether and when such methods truly are apt to investigation of the internal world.[6] In the background of such debates lies the central question of whether psychoanalysis is a science. If so, of what sort, and are our existing conceptions of science adequate when it comes to the idea of a 'science of subjectivity'? If not, is this because psychoanalysis is unscientific (i.e. a failed science) or non-scientific (e.g. a *Weltanschauung*)? More recently one thinks of the theory and findings of experimental psychology concerning the non-dynamic unconscious and their significance for psychoanalytic theory. In defending the claim that apparently meaningless human phenomena may have a sense, Freud (1916: 251) argued that merely physical (e.g. genetic or neurological) explanations frequently fail to tell us all we want to know, and that psychological explanation remains called for—but he failed to adequately consider the different psychological ways in which we may make sense of such phenomena (e.g. via such heuristics and biases in information processing as form part of an 'adaptive' unconscious). Recent work on both sides explores the complementarity of such explanations, and offers us the understanding that cognitive processing is motivated in psychodynamic ways and that the psychodynamic unconscious may be comprised of structures first delineated outside psychoanalytic theory (Eagle 2013; Chen and Chaiken

1999). Here the philosopher's role is both synoptic (surveying the points of overlap and contact in the objects and the theories) and diplomatic (working to ensure that different schools, with their different approaches to the life of the mind, do not talk past one another).

A certain kind of good psychoanalysis might go something like this (but without the linear form): build enough trust between a vulnerable patient and a respectful analyst; examine and carefully deconstruct the patient's defensive character formations; try to tolerate, truthfully acknowledge, and integrate such latent unintegrated and undeveloped feelings and expectations that induce shame and distrust in the patient; facilitate thereby the development of these feelings and the patient's increased realistic self-confidence. Along the way such grandiose ambitions and self-deceiving illusions as serve defensive ends may be dismantled in the pursuit of: a more workable inner life, an increased ability to remain inwardly and outwardly truthful, and the forming of deeper relationships. A good philosophical analysis of psychoanalysis may proceed along parallel lines. Having one's precious psychoanalytic understandings subjected to philosophical critique may be galling, parts of what was cherished may have to be foregone, ambitions may sometimes need to be scaled back, collaborations more willingly entered into—with the rewards being greater clarity and the opportunity for what is truly valuable within the theory to shine and grow. The result is a discipline with its finger even more keenly on the pulse of our baffling inner lives and yet more serviceable to those seeking to follow the Delphic command.

Truth and Truthfulness

We close with an illustration of the closeness of matters philosophical and psychoanalytic in relation to the Delphic command by an examination of their mutual concern with truth and truthfulness.

Freud commented that 'Psychoanalytic treatment is founded on truthfulness' (1915: 164) and—perhaps less truthfully!—that, regarding his own development of psychoanalysis, 'My single motive was the love of truth' (quoted in Sterba 1982: 115). It is a motivation he urges on the patient too, in the 'fundamental rule' of psychoanalysis (Thompson 2004). 'The only exception' to the rule that the patient be encouraged to speak of whatever he or she wants is 'in regard to the fundamental rule of psycho-analytic technique which the patient has to observe'. What is the rule? Freud (1913: 134–5) tells his patient that:

> You will notice that as you relate things various thoughts will occur to you which you would like to put aside on the ground of certain criticisms and objections. You will be tempted to say to yourself that this or that is irrelevant here or that it's quite unimportant or nonsensical so that there's no need to say it. You must never give in to these criticisms, but must say it in spite of them—indeed, you must say it precisely because you feel an aversion to doing so. Later on you will find out and

learn to understand the reason for this injunction, which is really the only one you
have to follow. So say whatever goes through your mind … [N]ever forget that you
have promised to be absolutely honest and never leave anything out because for
some reason or other it is unpleasant to tell it.

This valuation of truth-telling is also central to that most psychoanalytical of pre-Freudian
philosophers, Friederich Nietzsche, who, in his last work *The Antichrist*, wrote: 'Truth has
had to be fought for every step of the way, almost everything else dear to our hearts …
has had to be sacrificed for it. Greatness of soul is needed for it, the service of truth is the
hardest service' (Nietzsche 1895/1968: 50). In *The Gay Science*, he explains that the '"will
to truth" does *not* mean "I do not want to let myself be deceived" but—there is no
alternative—"I will not deceive, not even myself"; *and with that we stand on moral
ground*' (Nietzsche 1882/2001: 344).

Earlier we considered the notion of knowing thyself, and tried to do more justice to the
idea of such a gain in knowledge than could be done by reading it in terms of increasing
one's stock of information about oneself. In particular we stressed self-knowledge as
overcoming self-deception and self-alienation and as relating realistically with others.
Later we will also go on to discuss 'knowing one's own mind' in the sense of arriving at
non-vacillating resolve and determinacy of thought and will. So too, in considering truth,
we do well to attend to uses of the concept which take us beyond notions of mere
correctness. Thus not only a judgement expressed in a proposition, but also plumb lines,
hearts, desires, and lovers, may all be true. And if we are true *to* someone (including
ourselves—recall 'to thine own self be true'), then correct judgement also does not seem
to come into it. We may here recall Martin Heidegger's (1927/1962: §44) recovery of truth
as *alethia* from otherwise hegemonic conceptions of truth as *adequatio*—the former to do
with something's self-revelation or unimpeded unconcealment, the latter to do with one
thing's correct representation of something else (see Gipps, this volume). We may think,
too, of Ryle's (1949: 183–4) discussion of what he called 'avowals'—utterances such as:

'I want', 'I hope', 'I intend', 'I dislike', 'I am depressed', 'I wonder' etc— and of
how we may be tempted by their form to misconstrue all the sentences in which
they occur as self-descriptions. But in its primary employment 'I want …' is not
used to convey information, but to make a request or demand. It is no more meant
as a contribution to general knowledge than 'please'. … Nor, in their primary
employment, are 'I hate …' and 'I intend …' used for the purpose of telling the
hearer facts about the speaker … They are things said in detestation and
resolution and not things said to advance biographical knowledge about
detestations and resolutions.

The truth of such avowals is not a function of their expressing correct judgements that
one hates or intends, but rather of their being expressions of the hate and intention in
question. When we express ourselves truly, or again truthfully, we typically speak 'from',
not 'about', our thoughts and feelings, and do so without perverting their articulation.

Introduction: Know Thyself

What none of these philosophers considers, in their talk of truth as the auto-revelation of Being to us (Heidegger) or as the auto-revelation of the human heart and mind to itself and others (Nietzsche, Ryle), is the significance, including the ethical significance, of an interpersonal commitment to truthfulness for the very constitution of what is there to be revealed. To put it otherwise: we do well to avoid considering the value of truthfulness only in relation to the expression of *what already has determinate psychological shape*, and instead to acknowledge its even more fundamental role in our minds *becoming made up*, in various senses of that idiom. Such a focus is provided by Williams's (2002) discussion of the virtues of truth and truthfulness, the psychoanalytic resonances of which should shortly become clear.

Williams begins with the observation that many of our thoughts do not already clearly take the form of a belief as opposed to a desire or, say, a wish (2002: 82). For sure, sometimes we do have:

> very determinate dispositions to assert certain things. But in many other cases, it is not merely the case that we do not know what we believe (though this is of course often true), but that a given content has not come to be a belief at all. What makes it into a belief may be that we are asked about the matter or about the belief and then have to decide whether we are prepared to assert it or not. How can that be, if assertions are expressions of belief? The answer is that assertions ... often give others a reason to rely on what we say, either as a statement of how things are, or as an expression of how they seem to us. So ... I have to consider what I am prepared, sincerely and responsibly, to assert. I ask myself what I believe, and that is, in such a context, the same question. The question should not be understood, however, as simply one of what I already believe; in trying to answer it I do not simply review my dispositions but consider my reasons for taking a given content to be true, and this is a question of what *I am* to believe.

A subject may be sincere in that he may come out with what is on his mind at any moment, but unless there is some consistency between what he says from occasion to occasion it will be hard to treat what he says as expressive of anything that dignifies the description of 'belief' (Williams suggests the phrase 'propositional mood' as a more fitting alternative). He may at first be 'awash with many images, many excitements, merging fears and fantasies that dissolve into one another' (2002: 195).

What will help a subject firm up his thoughts into distinct beliefs, desires, and wishes that no longer bleed into one another, will be in part the conversations he has with others. In conversation you may ask me what I think or feel about something, and if I am to respect the relationship we have, to be of use to you, and to be someone whose word counts for something, it will be important that I give thought to the matter at hand and actually form determinate thoughts or feelings. At a level more basic than the enjoyment of any transparent self-understanding of determinate beliefs and desires 'we are all together in

our social activity of mutually stabilizing our declarations and moods and impulses into becoming such things as beliefs and relatively steady attitudes' (Williams 2002: 193).

Williams identifies a second way in which conversations clarify what we believe. Some of our thoughts are wishes, and through wishful thinking, turn into beliefs. Or again, some of our indeterminate thoughts may become either wishes or beliefs, and which they become may depend on other wishes and desires we have. Wishful thinking, says Williams 'is very basic and not a great mystery: the steps from its being pleasant to think of P, to its being pleasant to think that P, to thinking that P, cover no great psychological distance' (2002: 83). As a result, 'there is no mystery about the fact that … an agent may easily find himself committed to [the] content [of his wishes and beliefs] in the wrong mode' (2002: 198). However, this does not happen transparently. When beliefs arise in these ways, when they 'become hostage to desires and wishes, they do so only as the result of hidden and indirect processes, against which the disciplines of the virtues of truth are directed' (2002: 83). And this is something that conversations with others can help prevent.

This applies not only to questions of what to believe, but also when thinking about what to do. Since:

> individual deliberation … is inherently open to wishful thinking … it needs the virtues of truth as much as purely factual inquiries need them. [So] thinking about what one individual should do can usefully involve more than one person: we can think about what I should do. This is not just because you may have experience and knowledge which I lack, but because your wishes are not mine—possibly not in their content, certainly not in their effects. [We] help to sustain each other's sense of reality, both in stopping wishes' becoming beliefs when they should not, and also in helping some wishes rather than others to become desires

> (Williams 2002: 198).[7]

The same may also be said of a third question, self-interpretation: we may equally helpfully think together about who I am.

The implications of Williams's philosophical argument for both psychoanalytic theory and practice are clear. Yet, arguably, they are equally relevant to philosophical practice itself. Nietzsche and Wittgenstein both identify a similar role for the will and its influence on philosophical thought as Williams identifies for the wish here, and with it the significance of an ethic of truthfulness in philosophy. Wittgenstein's remarks return us to the very first sense of 'know thyself' we identified in relation to philosophy, the practical endeavour of the philosopher to come to know herself:

> What makes a subject difficult to understand—if it is significant, important—is not that some special instruction about abstruse things is necessary to understand it. Rather it is the contrast between the understanding of the subject and what most

people want to see. … What has to be overcome is not a difficulty of the intellect, but of the will. … Work on philosophy is … actually more of a kind of work on oneself

(Wittgenstein 2005: 86).

The edifice of your pride has to be dismantled. And that means frightful work. … One cannot speak the truth, if one has not yet conquered oneself. One cannot speak it – but not because one is still not clever enough

(Wittgenstein 1980: 30–5).

A Handbook of Philosophy *and* Psychoanalysis

Wittgenstein's reflections on philosophical practice carry both explicit and implicit psychoanalytic overtones (Baker 2003), but he famously had an ambivalent relationship to psychoanalysis itself (Levy, this volume), while Freud in turn famously had an ambivalent relationship with philosophy. In 1886 Freud avowed to Fliess that he secretly nourished 'the hope of arriving … at my initial goal of philosophy' after 'the detour of medical practice' (Freud 1985: 159). Furthermore he supplemented his medical lectures by enrolling in six of the philosopher Brentano's lecture courses, and also met with him outside the class (Tauber 2010: 29). He conceived of psychoanalysis as standing in a 'middle position between medicine and philosophy'; he also added that he had 'never really been a doctor in the proper sense' (cited by Rieff 1959: 301). Later though he confessed to a 'constitutional incapacity' for philosophy (Freud 1925: 60) and became famously dismissive of 'the philosophers'—whom he equated with those who dogmatically insisted that '"consciousness" and "mental" were identical' and so would not accept what he variously described as the 'postulate' or 'hypothesis' of unconscious mental life (1925: 31).

By happy contrast with Freud's 'philosophers', those contributing to this Handbook show a sympathetic interest in psychoanalysis' most central concept, the unconscious, in relation to its closest conceptual allies: defences, transference, conflict, free association, wish-fulfilment, and symbolism. Several of their chapters work to help psychoanalysis know itself by elucidating, retheorizing, and rescuing 'the unconscious' from objections and misrepresentations—including its self-misrepresentations. Other contributions explore psychoanalysis in relation to: its philosophical prehistory, the recognition and misrecognition afforded it within twentieth-century philosophers, its scientific strengths and weaknesses, its applications in aesthetics and politics, and its value and limitations when brought to bear on ethics, religion, and social life.

Further introduction we save for the openings of each of the sections which follow. Within each section, we have endeavoured to provide an evaluative overview of current thinking at the interface between philosophy and psychoanalysis through original contributions that will shape the future of the debate. Some chapters lean more towards the overview, others towards developing a line of argument that defends a particular position, but taken as a whole, each section forms the ground for future research.

We close by acknowledging three points. First, we have not sought to provide an introduction to psychoanalytic theory, for which we refer the reader to Bateman and Holmes (1995), Rusbridger and Budd (2005), Eagle (2011), Lear (2011), or Milton, Polmear, and Fabricius (2011). Second, there are some limitations in coverage. The reader will notice, for example, that Jungian, Lacanian, and feminist traditions feature far less prominently than Freudian and object-relations approaches. Sometimes this wasn't for lack of trying to solicit contributions but, especially in relation to the paucity of coverage of post-Lacanian developments, we must also own a lack of editorial expertise in assessing their cogency. The third relates to this being a Handbook of philosophy *and*, not philosophy *of*, psychoanalysis. It may be remarked that we have included some, but not extensive, coverage of the 'Freud Wars' which constituted a once-prominent strand within the philosophy of psychoanalysis of the last forty years. Some of the important issues that arose in those debates, in particular the scientific evidence for and conceptual validity of psychoanalytic theory, have naturally found their place in this Handbook, but on the whole, we have encouraged different modes of engagement of the kind discussed and illustrated in this introductory chapter.[8]

References

Baker, G. (2003). 'Wittgenstein's method and psychoanalysis'. In K. W. M. Fulford, K. Morris, J. Sadler, and G. Stanghellini (eds), *Nature and Narrative* (pp. 57–73). Oxford: Oxford University Press.

Bateman, A. and Holmes, J. (1995). *Introduction to Psychoanalysis*. London: Routledge.

Bettelheim, B. (1982). *Freud and Man's Soul*. New York: Knopf.

Carman, T. (2003). 'First Persons: On Richard Moran's *Authority and Estrangement*'. *Inquiry 46*: 395–408.

Chen, S. and Chaiken, S. (1999). 'The heuristic-systematic model in its broader context'. In S. Chaiken and Y. Trope (eds), *Dual process theories in social psychology* (pp. 73–96). New York: Guilford.

Eagle, M. (2011). *From Classical to Contemporary Psychoanalysis*. London: Routledge.

Eagle, M. (2013). 'The implications of conceptual critiques and empirical research on unconscious processes for psychoanalytic theory'. *Psychoanalytic Review 100*(6): 881–917.

Eissler, K. R. (1963). 'Notes on the psychoanalytic concept of cure'. *Psychoanalytic Study of the Child 18*: 424–463.

Freud, S. (1900). *The Interpretation of Dreams*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 5. London: Hogarth Press.

Freud, S. (1901). *The Psychopathology of Everyday Life*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 6. London: Hogarth Press.

Freud, S. (1913). 'On Beginning the Treatment'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 12 (pp. 121–144). London: Hogarth Press.

Freud, S. (1915). 'Observations on transference love. Further recommendations on the technique of psycho-analysis. III'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 12 (pp. 159–173). London: Hogarth Press.

Freud, S. (1916). *Introductory Lectures on Psychoanalysis*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vols 15 and 16. London: Hogarth Press.

Freud, S. (1925). *An Autobiographical Study*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 20. London: Hogarth Press.

Freud, S. (1985). *The Complete Letters of Sigmund Freud to Wilhem Fliess 1887–1904*, trans. J. M. Masson. Cambridge, MA: Harvard University Press.

Gaita, R. (2004). *Good and Evil: An Absolute Conception*, 2nd edn. Abingdon: Routledge.

Gomez, L. (2005). *The Freud Wars: An Introduction to the Philosophy of Psychoanalysis*. Hove: Routledge.

Hadot, P. (1995). *Philosophy as a Way of Life: Spiritual Exercises from Socrates to Foucault*. Oxford: Blackwell.

Hadot, P. (2002). *What is Ancient Philosophy?* Cambridge, MA: Belknap Harvard.

Hegel, G. W. F. (1807/1977). *Phenomenology of Spirit*, trans. A. V. Miller. Oxford: Oxford University Press.

Heidegger, M. (1927/1962). *Being and Time*, trans. J. Macquarrie and E. Robinson. Oxford: Basil Blackwell.

Jones, E. (1940). 'Sigmund Freud 1856–1939'. *International Journal of Psycho-Analysis 21*: 2–26.

Jones, E. (1955). *Sigmund Freud Life and Work, Volume Two: Years of Maturity 1901–1919*. London: Hogarth Press.

Kitcher, P. (1992). *Freud's Dream: A Complete Interdisciplinary Science of Mind*. Cambridge, MA: MIT.

Kohut, H. (1973). 'Psychoanalysis in a troubled world'. *Annual Psychoanalysis 1*: 3–25.

Lacewing, M. (2014). 'Psychodynamic psychotherapy, insight and therapeutic action'. *Clinical Psychology: Science and Practice 21*: 156–173.

Lear, J. (2005). 'The Socratic method and psychoanalysis'. In S. Ahbel-Rappe and R. Kamtekar (eds), *A Companion to Socrates* (ch. 27). Oxford: Blackwell.

Lear, J. (2011). *Freud*. London: Routledge.

Lear J. (2014). 'Integrating the non-rational soul'. *Proceedings of the Aristotelian Society 114*: 75–101.

Malcolm, N. (1958). *Ludwig Wittgenstein: A Memoir*. London: Oxford University Press.

McDonald, M. (2014). *Hegel and Psychoanalysis: A New Interpretation of 'Phenomenology of Spirit'*. New York: Routledge.

Menninger, K. A. (1956). 'Freud and American psychiatry'. *Journal of the American Psychoanalytic Association 4*: 614–625.

Milton, J., Polmear, C., and Fabricius, J. (2011). *A Short Introduction to Psychoanalysis*. London: Sage Publications.

Moran. R. (2001). *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton, NJ: Princeton University Press.

Nehamas, A. (1998). *The Art of Living: Socratic Reflections from Plato to Foucault*. Berkeley, CA: University of California Press.

Nietzsche, F. (1882/2001). *The Gay Science*, ed. Bernard Williams, trans. Josefine Nauckhoff. Cambridge: Cambridge University Press.

Nietzsche, F. (1887/1967). *On the Genealogy of Morals*, ed. and trans. Walter Kaufmann. New York: Vintage Books.

Nietzsche, F. (1895/1968). *The Antichrist*, ed. and trans. R. J. Hollingdale. New York: Viking Penguin.

Rieff, P. (1959) *Freud: The Mind of the Moralist*. New York: Viking Press.

Rusbridger, R. and Budd, S. (2005). *Introducing Psychoanalysis: Essential Themes and Topics*. London: Routledge.

Ryle, G. (1949). *The Concept of Mind*. London: Hutchinson.

Saussure, F. de (1915/1961). *Course in General Linguistics*, trans. W. Baskin. New York: McGraw-Hill Book Company.

Smith, David Livingstone (2003). *Psychoanalysis in Focus*. London: Sage Publications.

Sterba, R. F. (1969). 'The Psychoanalyst in a world of change'. *Psychoanalytic Quarterly 38*: 432–454.

Sterba, R. F. (1982). *Reminiscences of a Viennese Psychoanalyst*. Detroit, IL: Wayne State University Press.

Tauber, A. I. (2010). *Freud, the Reluctant Philosopher*. Princeton, NJ: Princeton University Press.

Taylor, C. (1985). 'Self-interpreting animals'. In C. Taylor, *Human Agency and Language: Philosophical Papers*, Vol. 1 (pp. 45–76). Cambridge: Cambridge University Press.

Thompson, M. G. (2004). *The Ethic of Honesty*. Amsterdam/New York: Rodopi.

Williams, B. (1995a). 'Saint-Just's Illusion'. In *Making Sense of Humanity* (pp. 135–150). Cambridge: Cambridge University Press.

Williams, B. (1995b). 'Nietzsche's Minimalist Moral Psychology'. In *Making Sense of Humanity* (pp. 65–78). Cambridge: Cambridge University Press.

Williams, B. (2000). 'Philosophy as a Humanistic Discipline'. *Philosophy 75*(4): 477–496.

Williams, B. (2002). *Truth and Truthfulness: An Essay in Genealogy*. Princeton, NJ: Princeton University Press.

Wittgenstein, L. (1980). *Culture and Value*, trans. P. Winch. Oxford: Basil Blackwell.

Wittgenstein, L. (2005). *Big Transcript*. Oxford: Blackwell.

Wollheim, R. (1984). *The Thread of Life*. Cambridge, MA: Harvard University Press.

Wollheim, R. (1993). *The Mind and Its Depths*. Cambridge, MA: Harvard University Press.

Notes:

(1.) Other important aims include self-becoming (i.e. developing what is as yet but latent within) and improved self-other relating. As will be evident from our discussion we take these to be interdependent with self-knowing. For elaborations see Eagle (2011) and Lacewing (2014).

(2.) The fantasy that one could avoid such philosophical troubles through avoiding philosophizing—and the correlative defensive equation of intellectual activity with intellectualizing defence—is akin to the fantasy that one may better go through life by simply avoiding troublesome emotional experience. The psychoanalyst, one might say, no more has the option of opting out of intellectual self-understanding than the philosopher has the option of opting out of the self-(mis)understandings that constitute our emotional lives. For self-(mis)understanding, in both emotional and intellectual registers, is an inexorable part of the human condition. One might say that one cannot not philosophize—only do so worse or better, i.e. only be less or more aware of the assumptions embedded in one's thinking.

(3.) In this way we can grasp why, say, a Woody Allen character who is preoccupied by psychoanalytically understanding himself is still neurotic: he is caught up in endless self-relationship rather than being emancipated from self-preoccupation.

It is important to note that more comes under the general concept of knowing thyself than the absence of unhelpful self-relations. Thus another important aspect of the concept refers to the cultivation of a straightforward determinacy of action. Such an agent takes responsibility for himself (he 'owns' his thought and his actions) and does not vacillate; he 'knows his own mind'. We return to this in the final section of this introduction.

(4.) A perspective most forcefully developed by Raimond Gaita (2004), and iterated by Backström (this volume).

(5.) This is but one of several ways in which Hegel prefigures psychoanalytic theory (McDonald 2014). We may think here too of Wittgenstein's plan to use either the Earl of Kent's 'I will teach you differences', or Joseph Butler's 'everything is what it is and not another thing', as the motto for his *Philosophical Investigations*. The thought is also familiar to us through Saussure's (1915: 120) consideration that 'in language there are only differences without positive terms': our concepts enjoy their determinacy in virtue of their exclusionary relations with other concepts, and we reflectively grasp what phenomena are by appreciating how they differ from other phenomena.

(6.) For summaries see Smith (2003) and Gomez (2005).

(7.) As Williams (2002: 198) notes, however, there can yet 'be a negative side to this same process: in helping you to decide, I may reinforce your fantasy, and we may conspire in projecting wishes into a deceptive social hologram'.

(8.) We are very grateful to Katy Abramson and Adam Leite for their comments on an earlier draft of this introduction.

Richard G.T. Gipps

Richard G.T. Gipps, Oxford University, Disability Advisory Service, Oxford, UK

Michael Lacewing

Michael Lacewing, Department of Philosophy, Heythrop College, University of London, London, UK

Oxford Handbooks Online

Abstract and Keywords

This section of the Handbook consists of five chapters that focus on how psychoanalysis intersects with the history of philosophy. Three themes are examined: philosophical anticipations of psychoanalytic ideas; the clarification of psychoanalytic ideas by situating them in their intellectual context; and alternative approaches to psychoanalytic material provided by philosophers such as Friedrich Nietzsche and Baruch Spinoza. Also considered in this section is how Nietzsche and Arthur Schopenhauer anticipated aspects of Sigmund Freud's psychoanalytic theory. The first chapter explores Schopenhauer's conception of mankind's motivations and his writings on madness, the second deals with Freud's thinking on sexuality and the sexual drive, and the third describes an implicit concept of an unconscious first made explicit by Friedrich Wilhelm Joseph von Schelling and later deployed to explicate human motivation by Schopenhauer and Nietzsche. The last two chapters discuss sublimation and the solipsistic aspect of Freud's systematization.

Keywords: philosophy, psychoanalysis, Friedrich Nietzsche, Arthur Schopenhauer, Sigmund Freud, madness, sexual drive, unconscious, human motivation, sublimation

Intellectual Prehistory: Introduction

As a distinctive therapeutic practice deploying free association as its key method, psychoanalysis may rightly be said to have originated with Freud. The same is not true of a variety of its core concepts such as the unconscious, repression, sublimation, projection, the significance of the sexual drives, and wish-fulfilling fantasy. The psychological and medical prehistory of psychoanalytic notions clustered around the idea of a dynamic unconscious has been masterfully traced by Ellenberger (1970). The chapters of this section consider instead the philosophical prehistory of psychoanalysis. They trace a history of ideas from Spinoza through to Nietzsche, and investigate three forms of connection between the two disciplines: 1) philosophical anticipations of psychoanalytic ideas; 2) the clarification of psychoanalytic ideas through situating them in their intellectual context; and 3) alternative—perhaps even superior—approaches to psychoanalytic material provided by the philosophers in question. While we have arranged this section's chapters in the chronological order of the philosophers they consider, in this introduction we instead discuss them in terms of the above three themes.

Philosophical Anticipations of Psychoanalytic Ideas

The two central philosophers who clearly anticipated aspects of Freud's (1856–1939) psychoanalytic theory are his precursor Arthur Schopenhauer (1788–1860) and his contemporary Friedrich Nietzsche (1844–1900). As described by Ellenberger (1970) and Chapman and Chapman-Santana (1995) amongst others, Nietzsche developed various concepts similar to those of Freud, for example: the transposition into psychology of the principles of the conservation and transformation of energy; sublimation; repression (*Hemmung*); 'ressentiment'; conscience theorized along the lines of the superego; unconscious guilt; and the id. In their chapter, Brook and Young provide a comprehensive elaboration of the ways in which Schopenhauer anticipated Freud. Central to Schopenhauer's conception of mankind's motivations in *The World as Will and Representation* (1819) is that of an inscrutable transcendental will, manifesting itself in sexual and self-preservative drives, with its own intentions that are yet disguised by the rationalizations of the ego. Brook and Young also draw our attention to Schopenhauer's striking writings on madness, comparing his consideration that, faced with unbearable sorrow, the mind 'seizes upon madness as the last means of saving life', with Freud's contention that delusions are 'applied like a patch over the place where originally a rent had appeared in the ego's relation to the external world'.

Comparison of Freud's psychoanalytic theory with that of Schopenhauer not only throws up interesting parallels but also, according to Sandford, helps reveal some of what is distinctively original about Freud's formulations. Sandford pursues this thought with especial reference to Freud's thinking on sexuality as found in the 1905 edition of his *Three Essays on Sexuality*. As Sandford describes it, Freud's predecessors, including Kant and Schopenhauer but also the sexologists Krafft-Ebing and Moll, held to a conception of the sexual instinct (*Geschlechtstrieb*) as geared towards copulation and reproduction. By contrast with this, Freud's conception of the sexual drive (*Sexualtrieb*) had it essentially aimed at the experience of pleasurable sensations, and only contingently connected to reproductive ends. As such, it provides us with a new philosophical anthropology and radicalizes our understanding of what can be meant by the concepts of the natural and the perverse in relation to human sexual life.

Psychoanalysis Situated in Its Intellectual Context

While Freud was critical of 'the philosophers', he appears to have had in mind the likes of his one-time teacher Brentano for the latter's exclusive focus on the conscious mind. Yet by contrast with Brentano, the German Romantics and idealists offered a conception of

the driving forces of the human subject as part of a will extending much further back into nature than is reached by consciousness (ffytche 2012). In his chapter, Gardner traces an implicit concept of an unconscious transcendental self underpinning consciousness, informing mental structure and creating mental content, back to Kant (1724–1804)[1] and Fichte (1762–1814). He tells how the concept was first made explicit by Schelling (1775–1854), and posited as generative of individual character by the physicians Schubert (1780–1860) and Carus (1789–1869), before being deployed to explicate human motivation by Schopenhauer and Nietzsche. Yet he also discusses how Freud's psychoanalysis makes use of another version of the self, one which Gardner exemplifies—without meaning thereby to attest to an informing influence—with the pre-Romantic Spinoza (1632–77). The aspect of Spinoza's philosophy to which Gardner attends are his naturalistic conception of affects as causally determining actions independently of the agent's rational endorsement. As explored throughout this Handbook, psychoanalysis is replete with fruitful yet challenging tensions between meaning, artistry, and understanding on the one hand, and causation, science, and explanation on the other. Gardner clarifies these tensions by describing how psychoanalysis contains within itself, in an imperfectly fitting manner, two different conceptions of mind: the experience-synthesizing transcendental mind of the post-Kantians, and the naturalistically conceived mind with causal powers and animal instincts of Spinoza.

Retheorizing the Unconscious

The third reason for considering psychoanalysis in relation to the history of philosophy is the resource the latter provides for providing alternative theorizations of psychoanalytic phenomena. This is most evidently undertaken by the twentieth-century philosophers considered in the next section, but also here by Gemes in his chapter on Nietzsche and sublimation, and by Macdonald in her chapter on Hegel and recognition.[2]

As Gemes describes, the concept of sublimation plays an essential role in psychoanalysis, yet was never given a fully satisfactory treatment by Freud. In particular, Freud failed to provide a convincing criterion to help us distinguish between symptoms caused by repression and culturally valuable activities resulting from sublimation. Nietzsche, however, now comes to the rescue on precisely this point by suggesting that, whereas neurotic symptoms involve disintegration of the psyche, sublimation involves the integration of sublimated drives within the master plan of a dominant drive. This dominant drive provides the lynchpin for our self-becoming, a notion which is essentially evaluative insofar as, by contrast with the ugly disintegrative and self-hating nature of repression, it involves power, greatness, beauty, and life affirmation.

As Gemes also makes clear, sublimation (by contrast with repression) involves a desire to unite, and in this way connects Nietzsche's thought with a conception of sublimation more prominent in post-Freudians such as Klein and Loewald than in Freud. Such a connection with post-Freudian themes is also made by Macdonald who draws on the dialectical conception of selfhood found in Hegel's *Phenomenology of Spirit* (1807) to help transcend what Ricoeur described as the solipsistic aspect of Freud's systematization (i.e. the structural, topological, and energy-balancing models of the individual mind) which clashes with the relational character of his therapy. Scholars influenced by Kojève's reading of Hegel have stressed the resources of the earlier Master–Slave moment of his dialectic (the struggle for recognition between, and the mutual dependency of, Slave and Master) for overcoming the merely intrapsychic character of Freud's structural and topological models of the psyche. This, however, inscribes a rather paranoid moment into the heart of human intersubjectivity, whereas Macdonald follows Hyppolite in suggesting a happier resolution. This interpretation focuses on later moments of the dialectic, in particular that of the Unhappy Consciousness and his relationship with the figure of the Counsellor. The resolution it brings helps the Unhappy Consciousness overcome his sense of meaninglessness and lack of self-sufficiency by grasping the intrinsically relational nature of any mature independence. This correlativity of mature independence and relational dependency has been placed centre stage by relational schools of psychoanalysis, yet receives a more rigorous dialectical formulation in Hegel's phenomenology.[3]

Clinical psychoanalysis is sometimes taken to task for its alleged aim of explaining to the patient his present pain in terms of what is uncovered by way of his history. Given the focus on the here-and-now transference relationship by many of today's psychoanalysts, the accusation can already be seen as baseless,[4] but the accusation also risks misunderstanding the value of such understanding of personal history as develops within an effective psychoanalytic treatment. One of the values of such historical work is the way it helps the patient grasp the non-inexorable contingency of her current moods and modes of relating (Lear, this volume). It may be clinically pointless to explain the present by relation to the putative past, but what is not pointless is a patient and analyst using reference to the past to help the patient get such a vantage on the present as reveals it to be optional, as one possibility amongst others, as something that can be conceived and rendered otherwise. Being able to step back and relate the present to the past is a restoration or new development of personal agency. We believe that the chapters offered in this section allow something of a similar shift to be made in relation to psychoanalytic theory. As psychoanalysis comes to know itself through narrating its own conceptual history, it is afforded the possibility of stepping back from and grasping its latent assumptions as contingent theories: as options that may both reveal and constrain, as developments that may be disputed and which could have been otherwise, as possibilities that now may indeed become otherwise in the hands of a reinvigorated theory-generating agency.[5]

References

Chapman, A. H. and Chapman-Santana, M. (1995). 'The influence of Nietzsche on Freud's ideas'. *British Journal of Psychiatry 166*: 251–253.

Ellenberger, H. F. (1970). *The Discovery of the Unconscious*. New York: Basic Books.

ffytche, M. (2012). *The Foundation of the Unconscious: Schelling, Freud and the Birth of the Modern Psyche*. Cambridge: Cambridge University Press.

Longuenesse, B. (2012). 'Kant's "I" in "I Ought To" and Freud's Superego'. *Proceedings of the Aristotelian Society, Supplementary Volume, 86*: 19–39.

Sandford, S. (2017). 'Freud, Bion and Kant: Epistemology and anthropology in *The Interpretation of Dreams*'. *International Journal of Psychoanalysis 98*: 91–110.

Notes:

(1.) See Longuenesse (2012) and Sandford (2017) for further investigation of the parallels between Kant, Freud, and Bion.

(2.) As Bernstein (this volume) also notes, Ricoeur also draws on Hegel in his reading of Freud.

(3.) See Lacewing (this volume) for further philosophical treatment of such issues in relational psychoanalysis as it often leaves these under-theorized.

(4.) See the section on 'Recognition versus Cognition' in Gipps (this volume) for elaboration.

(5.) We explore the various significances of self-knowledge for psychoanalysis in Gipps and Lacewing (this volume).

Michael Lacewing

Michael Lacewing, Department of Philosophy, Heythrop College, University of London, London, UK

Richard G.T. Gipps

Richard G.T. Gipps, Oxford University, Disability Advisory Service, Oxford, UK

Oxford Handbooks Online

Abstract and Keywords

This chapter sketches a reconstruction of the basic psychoanalytic conception of the mind in terms of two historical resources: the conception of the subject developed in post-Kantian idealism, and Spinoza's laws of the affects in Part Three of the Ethics. The former supplies the conceptual basis for the psychoanalytic notion of the unconscious, while the latter defines the type of psychological causality of psychoanalytic explanations. The imperfect fit between these two elements, it is argued, is reflected in familiar conceptual difficulties surrounding psychoanalytic theory and explanation.

Keywords: subject, causality, affect, unconscious, post-Kantian idealism, Spinoza

Sebastian Gardner

(p. 29) Introduction

INQUIRY in the history of philosophy may serve a variety of purposes and assume a variety of forms. In addition to tracking the actual historical story of the formation of ideas, we may seek to illuminate the thought of historical figures by drawing systematic comparisons in the absence of any relation of historical influence. Longuenesse's account (2012) of the relation between Kant's 'I ought' and Freud's superego fits this description, and I am in full agreement that there is insight to be had by drawing out systematic affinities between concepts, even when these are historically remote in the sense of belonging to quite different schools or traditions, or of participating in quite different intellectual projects, as is evidently the case with Kant and Freud.

Another thing which we may legitimately do, with the same end of philosophical illumination rather than strict historical explanation in view, is to offer systematic reconstructions of concepts and theories in terms of resources supplied by the history of philosophy. In this vein, I am going to use the space available to suggest a reconstruction

of psychoanalytic theory—more specifically, of the basic conception of the mind evidenced in Freud's metapsychology—in terms of historical materials. The reconstruction falls into two parts. The first proposes a source for the psychoanalytic conception of the unconscious in post-Kantian idealism. The second attempts to locate a point of affinity of psychoanalytic explanation with Spinoza. Although I regard the first as a defensible historical claim—I think it is plausible to maintain on historical grounds that the legacy of idealism transmits itself into Freud's model of the mind, and I will formulate my account in such terms—for present purposes it can be taken in a purely systematic spirit, which is in any case how my second claim, concerning Spinoza, is intended. What follows (p. 30) may be regarded, therefore, as a purely analytical exercise cast in the form of a *Just-So* story: in a nutshell, I am suggesting that psychoanalytic theory may be viewed as the product of combining a conception of the subject developed in post-Kantian thought with a Spinozist view of the mechanics of human motivation; as if Freud had come up with the metapsychology as a facilitating framework for developing empirical psychological hypotheses through a reading of classical German philosophy and the *Ethics*. In the final section of the paper I will suggest that the idealist and Spinozist dimensions of psychoanalytic theory do not cohere fully, and that the tension between them is responsible for and reflected in certain familiar conceptual difficulties that have been held to confront Freudian theory. The ultimate point of the reconstruction is thus to shed some light on the famously puzzling character of psychoanalytic theory and explanation, in a way which bears on Longuenesse's alignment of Kant with Freud.

The Unconscious in Post-Kantian Idealism

One component of psychoanalytic theory which is both fundamental to it and, as it seems to me, not likely to be well accounted for by locating Freud's metapsychology in a line of descent solely from developments in nineteenth-century scientific psychology, is the opposition which lies at its core between consciousness and the unconscious. This distinction is not simply equivalent to, even if it at some point involves, a distinction between mental contents characterized in terms of different degrees of accessibility to introspection or conscious reflection, or of different (for instance, personal versus sub-personal) modes of processing (different types of) information. Rather it invokes as its starting point the ordinary idea of the person as a self-conscious self-determining subject, to which it then counterposes the notion of an enduring mental structure, capable of changing its contents over time, causally disposed to maintain itself in existence, systemically distinct from the network of mental states manifest in the stream of consciousness, and governed by a heterogeneous set of laws: a structure which is nonetheless posited *within* the orbit of the subject and held to stand in an intimate, immediately determining relation to conscious and self-conscious mental life.

This idea is, as Freud appreciated, highly puzzling: common-sense psychology accommodates mental states which are in various ways obscure, subliminal, latent, implicit, and so on, but nothing in it prepares for the notion that, as it has been put, I am (also) my Other. Freud, of course, spells out several justifications for the concept of the

unconscious, but these are best regarded as essentially warrants for employing an implicitly pregiven concept, not *ab initio* explications of a novel theoretical construction. So we have the question: how, through what sequence of reflections, can it come to be thought that persons as ordinarily conceived harbour counterselves, abiding configurations organized independently from their self-conscious subjectivity, their relation to which (p. 31) naturally draws comparison with a relation to another person, yet with which they are required to identify themselves?[1]

At least the rudiments of this idea can be arrived at through epistemological reflection of the type pursued by Kant's immediate successors.[2] The usual suspects cited when nineteenth-century antecedents of psychoanalysis are sought include Schelling, Romantic *Naturphilosophen* such as Gotthilf Heinrich von Schubert, Ignaz Paul Vital Troxler, and Gustav Carus, Eduard von Hartmann, and, famously, Schopenhauer and Nietzsche. In all of these thinkers an identification is made of the unconscious portion of the subject with Nature and the forces holding sway in it, which is also, of course, a foundational tenet of psychoanalysis; but in order for that identification to make sense, the concept of the former had to be available initially, and in order to locate that original notion, we need to look further back, to the first generation of thinkers occupied with the post-Kantian transcendental project.

The ground for the development of a transcendental unconscious was, in effect, prepared by Kant himself: in the first place, through his notion (as the first *Critique* allows itself to be and has been read, whether or not correctly) that mental contents are not conscious per se, in and of themselves, but need to be *made* conscious, and that cognitive awareness of both self and objects involves conceptually motivated operations, comprising synthetic acts not given in inner sense; and second, through his thesis of the necessary limit of self-knowledge which is due, not to the impossibility of theoretically cognizing noumena, but to the consideration that any attempt to cognize the self qua subject of thought 'can only revolve in a perpetual circle' (Kant 1781/1999: A346/B404), it being 'very evident that I cannot know as an object that which I must presuppose in order to know any object' (Kant 1781/1999: A402).

The notion that self-consciousness contains an aporia—that the 'I or he or it (the thing) which thinks' must represent itself as and only as '=X', a something 'known only through the thoughts which are its predicates' (Kant 1781/1999: A346/B404)—undergoes a radical modification in Fichte's *Wissenschaftslehre*. The complex theory of mental acts which Fichte provides in its Theoretical Part shares with Kant's theory of synthesis the aim of establishing the preconditions of empirical consciousness, but has a very different character. Kant's agencies and acts of synthesis, though non-conscious, can still be said to manifest themselves directly in conscious mental life, insofar as Kantian syntheses are necessarily discernible in—can be read off from—the conceptually shaped (p. 32) sensible given. By contrast, Fichte's multilayered account of the absolute I's positing and counter-positing, in which elements are cross-related in oppositional structures, identified without reference to recognizable types of mental representation, and form genetic series in which successive forms displace their predecessors, are set at a vastly greater distance

from the conscious surface of mental life. Fichte's speculative structures do not merely shadow what is present in natural consciousness but seek to bring to light its complete sufficient ground, hence their greater depth and remoteness.

What in particular screens Fichtean transcendental grounds from ordinary consciousness is the operation of the faculty which he calls productive imagination. The transition to empirical consciousness involves, on his account, an abrupt discontinuity, which has no conceptual representation and so must be effected by imagination, which converts purely intelligible structure into the kind of (ultimately, spatio-temporal) mental content that an empirical subject can recognize as its own (Fichte 1794/1982: 187–8, 193–4, 200–1). Experience depends upon an interplay and 'clash' of opposed components, but this is possible, Fichte argues, only if the 'boundary' between them is positively represented, which is something that no purely intellectual function can do: the role of productive imagination is accordingly to give phenomenologically concrete, intuitable form to a structure which pure thought can grasp only as a relation of irreconcilable mutual exclusion. Fichte draws an analogy with the way in which instants of light and darkness can be given in experience as alternating only if the boundary between them is extended into a temporal instant (Fichte 1794/1982: 187–208): in similar fashion, the I can grasp itself as distinct from its objects, and figure as an object for itself, only by relating to itself through the medium of imagination.

The upshot of Fichte's account, as regards the theme that we are tracking, is complex. In one sense, the *Wissenschaftslehre* eliminates the aporia that Kant locates in self-consciousness: though the I qua subject cannot be given objectually within empirical self-consciousness, it is not unknowable, since philosophical insight into the self-positing I provides knowledge of the thing that for Kant remains a mere '=X'. In another regard, however, Fichte reaffirms and underlines the aporia: the genesis of empirical self-consciousness from absolute I-hood covers its tracks, and the standpoints of life and philosophy are distinct and exclusive, the way things appear from the one being an inversion of how they appear from the other (Fichte 1794/1982: 207–8).

Although, for Fichte, the transcendental grounds of ordinary mental life are invisible to the pre-philosophical subject, the relation between them is open to full philosophical comprehension. Such epistemological optimism is open to challenge, however. Two figures are of particular importance in the present context. Kant's early critic Salomon Maimon (1790/2010) argued that Kant's theory of empirical knowledge fails to close the gap between the a priori conditions which issue from his transcendental proofs and the a posteriori contents of empirical judgement, to the advantage of Hume. In order to address scepticism, it is necessary, Maimon argued, to embrace a more rationalistic form of idealism than Kant's. Maimon proposes, accordingly, a Leibnizian recasting of the Kantian divisions between spontaneity and receptivity, and apriority and aposteriority, in which what he calls 'infinitesimals of sensation', akin to *petites perceptions* and (p. 33) belonging to the subject qua passive, are postulated as the elements of synthesis, which proceeds without consciousness and in accordance with rules not given to our understanding. This gives rise to the *appearance* of an a posteriori sensible given. In a

cognitively perfected consciousness, Maimon supposes, Kant's divisions would be overcome, and such a subject would grasp itself and its power of cognition as parts of an infinite mind or reason. On Maimon's account, the ground of sensation is located therefore within the subject, and derives ultimately from the deficiency of self-understanding which is constitutive of ordinary consciousness.

The same broad type of development—this time in reaction to Fichte—is found in Novalis, who provides a clear exposition of the shared view of Romantic post-Kantians that the ground of consciousness is necessarily inaccessible to discursive reflection and so must be grasped by other, non-discursive means. Novalis rejects Fichte's claim that the concept of positing can lead us to the absolute ground of self-consciousness and empirical cognition, on the basis that the structures of opposition and non-identity which positing involves are themselves products of, and valid only within, our limited cognitive structure (Novalis 1796/2003: 167–8). The consequent problem of elucidating being prior to our representation of it in predicative and identity statements is then provided with an affective-aesthetic solution: if being qua ground of consciousness must be considered in abstraction from any objectual character, then its proximate manifestation within us must be similarly non-objectual, that is, it must comprise a mode of consciousness which is not intentionally directed, which is as much as to say that it must have the character of feeling. Feeling must, however, stand in some relation to our judgementally articulated consciousness in order for it to function as a transcendental ground. This relation, Novalis argues, is one of content to form, but it follows what Novalis calls the principle of 'ordo inversus': reflection reverses the true relations obtaining within the subject, on the analogy with a mirror image, so that when reflection takes up feeling, being's affective self-manifestation is lost from view, and the status of being *something* is attributed instead to what has been formed conceptually.

This provides a rough indication of how the post-Kantian model of the subject can evolve, under the momentum of transcendental reflection, in a psychoanalytic direction. The transcendental project, as Dieter Henrich has argued (2003: ch. 2), aims to construct a theory of the subject in which the subject can recognize itself: it specifies the conception under which the mind operates and which is deployed implicitly in making those operations possible. This subjective fit provides one measure of validation for the theory. The transcendental image of the mind should thus be or correspond to—in a phrase of Richard Wollheim's (1972)—the mind's own image of itself. This conception of the end and method of transcendental inquiry does not, however, as we have seen, bind it to representing the subject as fully self-transparent: for a variety of reasons, gaps may be affirmed in the transcendental theory of the mind, and where these exist they will imply regions of opacity in the subject's apprehension of its own grounds.

The next historical and logical step consists, as indicated earlier, in the insertion of Nature into the place opened up by post-Kantian transcendental reflection. The conception of the empirical subject as constituted by an aporetic self-relation—the mind as a (p. 34) dynamically self-concealing structure, which can open up to itself and the world only on the condition that the facilitating ground of its doing so is excluded from

consciousness—is annexed to the philosophy of nature by Schelling (1800/1978), who first formulates explicitly the concept of the Unconscious, identifying it with the infinite activity of *natura naturans*. Schelling's unconscious has no specific role in psychological explanation, but Schubert, Carus, and others in his wake ascribe an unconscious to the human subject in an individualized form: the development of human personality is regarded as the effect of a subsisting unconscious core, which expresses in microcosmic form the principles, forces, and order of Nature at large. The final chapters in the story of progressive approximation to the psychoanalytic unconscious are familiar: Schopenhauer and then Nietzsche advance speculative hypotheses concerning the foundations of human motivation which repudiate the original transcendentalist association of the unconscious with rationality.

Spinoza on Human Motivation

In Part Three of the *Ethics*, 'Concerning the Origin and Nature of the Emotions', Spinoza devotes many pages to a systematic catalogue of the psychological laws governing the affects. On the basis of three 'primary' emotions of pleasure, pain, and desire, and their immediate derivatives, love and hatred (E3 P11, E3 P13 S; 2002: 285, 286), Spinoza defines in Propositions 13–59 (2002: 286–311) the complex kinds of emotion which may be formed, and the conditions under which they increase or decrease in intensity, and yield to, or are transformed into, one another.

Spinoza's purpose in giving this account is, at a minimum, to bring home to us the ways in which, and the degree to which, our passions determine our action dispositions independently of anything that could be called reflective endorsement. The account thereby serves the end of increasing our freedom: correct insight into the causation of our passions—above all, appreciation that they cannot be expected to respond to reason in the way that we naively suppose—can spur us to adopt the less direct but truly effective measures for their regulation described by Spinoza in the opening pages of Part Five.

As commentators have observed, there is much in the *Ethics* that can be regarded as proto-psychoanalytic. Spinoza envisages explanation of mental states and changes in a way that parallels and tracks bodily conditions and changes. Spinoza's strict psychological determinism echoes Freud's equally categorical statements on the topic. The picture of the mind as moving from one condition to another according to laws of efficient causality—registered by changes of hedonic state but not explained by invoking pleasure as an end—recalls the model that Freud draws up in his pre-psychoanalytic *Project for a Scientific Psychology* and which survives recognizably in the 'economic' component of the psychoanalytic metapsychology. The concept of imagination, given very broad scope in the *Ethics* and introduced primarily in order to account for the limitations of human cognition, is employed by Spinoza in the context of the affects in ways that show (p. 35) it doing duty for the panoply of concepts (pleasure as opposed to reality principle, primary as opposed to secondary process, phantasy, screen memory, etc.)

employed by Freud to explain the mind's capacity for operating with indifference to reality. Spinoza, like Freud, regards wish-fulfilment as having its basis in the fundamental structure of the mind.[3] Spinoza's assertion of the affective impotence of cognitions, their inability to check emotions directly by virtue of their truth alone,[4] is mirrored in Freud's account of the insufficiency of intellectual assent to interpretations and of the necessity of working-through. Spinoza anticipates Freud's devaluation of the property of consciousness, if not in fact to have taken it a good deal further, insofar as consciousness barely figures in the *Ethics*, which nowhere treats it as a topic in its own right. Conatus corresponds to Freud's conception of libido. And so on.

What I want to focus on, with a view to getting a unified perspective on these doctrinal points of contact of Spinoza with Freud, is the *nature* of Spinoza's claims in E3 P13–59.

Some of the propositions advanced in E3 P13–59 are indisputably true, perhaps even, as we would put it, conceptual truths, while others are far from obvious and require decoding. Spinoza gives no explicit account of their epistemological basis, and their status is not immediately clear. Insofar as they deal with observable patterns in the phenomena of psychological life, it would be natural to suppose that those which are not candidates for conceptual truth must be a posteriori, but it is barely conceivable in light of the method and ambition of the work as a whole, and the specific role of Part Three, that Spinoza should intend a mere inventory of psychological tendencies drawn from common observation, amounting to the inferior kind of knowledge that he calls 'casual experience' (E2 P40 S2; 2002: 267). Spinoza's aim must be, rather, to nail down the essential 'logic' of affect—in the context of the *Ethics*, nothing less would do.

This raises the expectation of being able to regard Spinoza's propositions as a priori, and Spinoza does indeed spell out, in the proofs of the propositions, deductive connections with the foundational metaphysical claims about human beings that he has made in Part Two and the earlier portion of Part Three. In these proofs Spinoza's constant reference point is his thesis of the mind's necessary endeavour '*to think of those things that increase or assist the body's power of activity*' or (what is the same) '*the power of thought of our mind*' (E3 P11–13; 2002: 284–6).

What deserves attention is the nature of the necessity whereby conatus expresses itself in the causation of affect. One type of necessity that we can make good sense of in the analysis of psychological explanation is the normative necessity associated with rational choice theory and the ordinary practical syllogism,[5] and some of Spinoza's propositions may look as if they can be expressed in such familiar terms: that he '*who hates someone will endeavour to injure him unless he fears that he will suffer a greater injury in return*' (p. 36) (E3 P39; 2002: 298), for instance, might be considered a prudential norm. But this cannot be what Spinoza has in mind, since such rational necessity is manifestly absent from his claims, for example, that we shall love or hate an object which we imagine to be similar to an object that is wont to affect us with pleasure or pain (E3 P16; 2002: 287); that we will share the emotion of whatever we imagine to be like ourselves (E3 P27; 2002: 292); that we love, desire, or hate a thing the more steadfastly when we think it

loved, desired, or hated by somebody else (E3 P31; 2002: 294); that we will seek to deprive someone of anything they enjoy that only one person can possess (E3 P32; 2002: 295); that our love for someone will turn to hatred if they form a more intimate bond of friendship (E3 P35; 2002: 296); that we will desire to possess a thing that gave us pleasure in the same circumstances as when we first took pleasure in it (E3 P36; 2002: 297); that we will hate someone whom we imagine to hate us without cause (E3 P40; 2002: 299) or someone imagined to be similar to ourselves who hates something that we love and imagine to be similar to ourselves (E3 P45; 2002: 301); that someone's hatred of an object that he has previously loved will be proportionate to the strength of his former love, and vice versa (E3 P38, E3 P44; 2002: 298, 301); that we will have higher regard for an object that we have not seen in conjunction with other objects or that has nothing in common with them (E3 P52; 2002: 304). The vectors and tendencies of emotional life to which Spinoza is pointing are familiar and, as we would put it, thoroughly *natural*, but if we choose to honour them on that count with some sort of claim to some kind of rationality, it is simply because of their familiarity, prevalence, and embeddedness in human nature, not because we suppose any articulable principle of reasoning or judgement of 'rightness' to underlie them.[6]

On a weak reading, of the sort that I offered earlier as a minimal characterization of his intentions, Spinoza would be taken to have shown only how the affects comport themselves when left to their own devices, information which we could regard as giving helpful forewarning of the ways in which our rational self-determination is prone to being subverted. This, however, would be to picture the affects as situated on the outside, and ourselves—as centres of rational self-determination—on the inside, in the situation of a swimmer calculating the forces of tide and current. A compelling reason for thinking that this view of our situation cannot be Spinoza's is that it corresponds to that of Descartes, as Spinoza construes—and explicitly rejects—it; it is to 'conceive man in Nature as a kingdom within a kingdom' (E3 Preface; 2002: 277).[7] That a much stronger reading is in order becomes clear when Spinoza says that 'mental decisions are nothing more than the appetites themselves', 'that mental decision on the one hand, and the appetite and physical state of the body on the other hand, are … one and the same thing' (E3 P2 S; 2002: 281),[8] a claim which coheres with his declaration of intent to 'consider human actions and appetites just as if it were an investigation into lines, planes, or bodies' (p. 37) (E3 Preface; 2002: 278). Spinoza envisages, therefore, the affects and their causation as occupying the full space of psychological explanation: they are not what we qua rational beings are set *over and against*, rather they are what we *are*, whence their *metaphysical* necessity;[9] the question of why we find ourselves subject to immediate affective determination receives its answer alongside the question of why we conceive physical bodies as divisible and numerable. Spinoza's logic of affect is designed to show that psychological life proceeds in ways that are not driven by any power of practical reason: explanations with the form of practical syllogisms do not express falsehoods—human action is not 'blind' in that sense—but they do not display the causal reality of the mind, since what makes them true, insofar as they are true, is not a grasping of normative relations. Recognition of the true, non-normative character of psychological

reality is the first step in escaping from human bondage: only when we have grasped that our 'ought' thoughts are nothing but further expressions of conatus, without *sui generis* reality, are we in a position to steer ourselves in the appropriate therapeutic direction.[10] There is therefore a good sense in which the ordinary concept of a reason for action does not survive the *Ethics*.[11] Spinoza's interchangeable use of 'reason' and 'cause' gives notice of this.[12]

What should accordingly strike us when we consider E3 P13–59 with Freud in mind is the remarkable parallel of Spinoza's explanations by and of emotion with psychoanalytic explanations, which similarly appear—from the standpoint of common-sense psychology—to be poised awkwardly on the border separating reasons from mere causes. Just as it would make no sense to cite the inherent tendency of love to convert itself into hatred in the circumstances detailed in E3 P35 as a *reason* for regarding a person with hatred, so similarly unconscious determinants as such cannot be incorporated into our self-reflection as justifications for our actions or propositional attitudes. And yet, nor are these forms of causation correctly described as *merely mechanical*: the relations between cause and effect are *internal*, and the effects are *intelligible* in light of their postulated causes.[13] There is no space here to go into detail, but I think it would not be hard to show, with reference to Freud's own texts and to typical psychoanalytic case histories, that the explanatory baseline in psychoanalytic explanation consists in the interaction of elements in psychic reality in ways which fuse (in a way that at once illuminates and puzzles) activity of thought with forms of activity conceptualized in terms of movements of organic and even inorganic bodies.

(p. 38) The shared nature of Spinozistic and psychoanalytic explanation—their fusion of thought and mechanism into a single form of efficient causality—is reflected in the points of doctrinal covergence listed earlier. Other revisionary claims concerning the human subject defended in the *Ethics* with conspicuous psychoanalytic resonance include Spinoza's reductive conception of the self, his rejection of free will, his use of the concept of expression as a linchpin of psychological explanation, his overhauling of the common-sense taxonomy of mental states, and his argument that the very fact of failures of reason suffices to falsify the kingdom-within-a-kingdom conception of the subject.[14]

The Tension in Psychoanalytic Theory

My answer to the question posed earlier—whence the psychoanalytic concept of the unconscious?—is therefore that Spinoza's logic of affect can be thought to mesh with the idealist conception of self-concealing subjectivity, yielding a model of the mind which approximates to that of psychoanalysis: in rough outline, the psychoanalytic conception of the unconscious is the product of substituting for the synthetic acts of transcendental subjectivity, postulated by the idealist/transcendentalist with a view to solving problems of epistemology and metaphysics, the sorts of contents and operations which Spinoza postulates with a view to explaining human action and affect. Freud has accordingly a foot in two distinct camps: the humanistic tradition of classical German philosophy, in

which cognitive and ethical tasks are conceived and pursued from the internal standpoint of the subject, and the alternative naturalistic programme of modern philosophy.

This seems to me a fair representation of the intellectual forces which come together in Freud's notion of the unconscious, and it accords well with the salient facts of psychoanalysis' historical reception—to wit, the huge appeal which psychoanalytic ideas have had across the board in the humanities for over a century, alongside the determined attempt, to the fore in the North American psychoanalytic tradition, to hold psychoanalysis fast (for better or worse) to the agenda of the natural sciences.[15] It is difficult to see that any more economical hypothesis could account for the extraordinary depth of disagreement separating different schools of psychoanalytic thought, ranging as it does from deconstructionist and neo-Nietzschean appropriations of Freud to the interpretation of psychoanalysis as proto-neurophysiology.

Representing psychoanalysis as an attempt to unite in a single concept two major and distinct trajectories in modern philosophy throws light on the theory and accounts for (p. 39) the double character of Freud's notion of the unconscious, but by the same token makes clear the difficulty which it faces. On the one hand, the legacy of idealism showing itself, Freud asks that we *avow* the unconscious, that we regard it as *our own*, while on the other he insists, in line with Spinoza, that the real forms of motivational causation are incongruent with the image of action as proceeding from a centre of self-conscious self-determination. This combination of claims is what forces out the characterization of the unconscious as the 'Other' of consciousness. The extended debate in the philosophy of psychoanalysis concerning whether psychoanalysis is a form of causal explanation or of reason explanation, reflects this duality.[16] So too does the puzzle, which has also given rise to an inconclusive literature, concerning the coherence of supposing that bona fide mental states, of the sort attributed in psychoanalytic explanation, can subsist without consciousness.

The problem can be put in sharper focus if we contrast the relation to common-sense psychology required for psychoanalytic explanation with the situation of Spinoza. As my earlier discussion emphasized, Spinoza's intentions are uncompromisingly *revisionary*. If Spinoza is not an eliminativist regarding common-sense psychology, he comes within a hair's breadth of being one. Psychoanalytic theory cannot, however, mean to follow Spinoza's hard line, because its own distinctive forms of explanation are interwoven with those of common-sense psychology, and its own concept of the unconscious is formulated with reference to and as the counterpart of the ordinary conception of the subject.

The consequences of this conceptual compromise—psychoanalysis' adoption of Spinozistic explanation, without acceptance of its revisionary implications—are reflected in Sartre's critique of Freud, the nub of which is contained in the following passage:

> By rejecting the conscious unity of the psyche, Freud is obliged to imply
> everywhere a magic unity linking distant phenomena across obstacles, just as
> sympathetic magic unites the spellbound person and the wax image fashioned in
> his like. The unconscious drive (*Trieb*) through magic is endowed with the

character 'repressed' or 'condemned', which completely pervades it, colours it, and magically provokes its symbolism. Similarly the conscious phenomenon is entirely coloured by its symbolic meaning although it can not apprehend this meaning by itself in clear consciousness.

(Sartre 1943/1958: 53–4)

This claim is explained more fully in Sartre's essay on the emotions:

The psychoanalytic interpretation [of an emotion] conceives the conscious phenomenon as the symbolic realization of a desire repressed by the censor. Note that, for consciousness, the desire is *not involved in its symbolic realisation*. In so far as it (p. 40) exists by and in our consciousness it is only what it gives itself out to be: emotion, desire for sleep, theft, laurel-phobia, etc. If it were otherwise, if we had any consciousness, even *only implicit*, of the real desire, we should be in bad faith, and that is not what the psychoanalyst means.

(Sartre 1939/2004: 30)

So the relation between an emotion and the signification conferred on it under a psychoanalytic interpretation must, to some degree, resemble the relation of the ashes of a fire extinct upon a mountain to those who lit it. And yet, Sartre points out, it cannot be *merely* external. For the psychoanalyst 'there is always an internal analogy between the conscious fact and the desire it expresses, since the *conscious fact is symbolical of the expressed complex*': the symbolic character of the conscious fact 'is *constitutive* of it'. This, however, leads us back in a circle, to the identification of unconscious motivation with mere bad faith: 'if symbolization is constitutive it is legitimate to see an immanent bond of *comprehension* between the symbolization and the symbol'—'the relation between symbol, symbolized and symbolization is an intra-structural bond of consciousness' (Sartre 1939/2004: 32).

The problem that Sartre describes, expressed in terms of my historical reconstruction, is therefore that psychoanalytic explanation invokes the conception of the unity of the mind and of the nature of relations between mental elements contained in our commonsense self-conception, elaborated in the Kantian tradition, *and also* models the mind in terms of Spinoza's patterns of causation, from which the same degree and type of reflexive unity is absent. The psychoanalytic concept of symbolization is a bridge, not just from unconscious to conscious facts, but also from the one philosophical model of the subject to the other, and the strain which it is under reflects the way in which psychoanalytic theory merely *superimposes*—grafts onto one another, without integrating—the two different models of the mind.[17]

There are, to be sure, ways of softening the impact of Sartre's criticism. We may, for instance, postulate some species of mental connection midway between immanent bonds of comprehension and merely external causal linkages of the fire and ashes sort. And it is clear furthermore that in doing so we may claim a degree of support from common-sense

psychology (which does not, after all, assimilate all 'bonds of comprehension' to the clear-cut case of knowing the meaning of a sign, as Sartre perhaps tends to imply). Moves of this sort can be given plausibility through being interrelated with one another in a systematic way. Psychoanalysis can thus be presented as an extension of common-sense psychology. What I think nonetheless needs to be recognized as the substantial truth in Sartre's criticism—and which in no way depends upon his professed Cartesianism—is his insight that, to the extent that we extend common-sense psychology by proposing a new species of mental state or relation, even though each individual theoretical move in the construction of the extension may have an adequate local justification, there is a sense in which the cumulative effect of the incremental changes is a *philosophical* loss of (p. 41) bearings; in historical shorthand, we wonder how, starting with Kant, we could have ended up with Spinoza. This is quite compatible with the resulting theory's working perfectly well as an explanatory system in accordance with criteria appropriate to empirical theories; what it means is only that we cannot render the theory philosophically perspicuous.[18]

One way in which the problem might be resolved decisively is through an expansion of the boundaries of the self, in a way that permits reconception of the unconscious in terms that make full metaphysical sense of the identification with it demanded by psychoanalysis. If the Nature that feeds into my unconscious is in reality more of myself, if it is a whole with which I am in some sense identical, then it is false that, as Sartre puts it, 'I *am* ego but I am not *id*' (1943/1958: 50): rather, the unconscious is a *continuation* of me. This view of the metaphysics of human personality is, in fact, the picture proposed (for epistemological and metaphysical reasons) by Schelling. It amounts, of course, to a revision of the common-sense view of the borders of the self, but it conserves, unlike Spinoza's metaphysics, the notion of self-conscious self-determination. However, it also, plainly, requires attributing to Nature properties that disenchanted modern natural science denies it, and though a hint of a preparedness to retrieve nineteenth-century Romantic idealism can be found in a very few psychoanalytic thinkers[19]—and arguably receives encouragement from Freud's late remarks on Eros and Thanatos—it is in flat contradiction with Freud's official account of the psychoanalytic *Weltanschauung* and not likely to be regarded as an acceptable condition for securing the conceptual integrity of psychoanalysis.

Where does all this leave psychoanalysis, as regards its systematic evaluation? The larger issues which arise concerning the long-running debate concerning psychoanalysis' claim to truth obviously cannot be pursued here, but I have a limited suggestion, arising from the preceding discussion.

One reaction to my account—which will recommend itself to those unimpressed by psychoanalysis' claim to psychological insight—is to take the diagnosis implied by the historical reconstruction as the basis for a quick conviction. The empirical defensibility of psychoanalytic claims is a matter for another occasion, but if the roots of (p. 42) psychoanalysis' conceptual difficulties lie as deep as I have suggested, namely in a collision of transcendentalist with naturalist commitments, then it may reasonably be

held that the re-emergence of this overarching philosophical antinomy in the psychoanalytic context—albeit in an especially overt, acute, aggravated form—cannot be a reason for impugning the theory's claim to offer genuine psychological explanation. Psychoanalysis ought not to be made a casualty of our inability to resolve what is at the end of the day a wholly general philosophical problem. Acknowledging the ambiguity concerning the personal versus Spinozistic status of the unconscious, and accepting that the theory does not explicate itself in a fully coherent manner, is an acceptable price to pay for a picture of human beings that makes sense of stretches of human experience which will otherwise be left completely opaque.

In so far as Longuenesse and I differ, our difference lies in our respective attitudes to the proposed unification of Kant's 'I ought' and Freud's superego. This Longuenesse regards as an unproblematic theoretical unification, which promises to help out Kant's ethical theory, by transposing a questionable transcendentalist conception into a secure naturalistic context. In my view, waiving all questions about the capacity of Kant's conception of morality to survive the move intact, the proposal draws attention to an underlying bifurcation in psychoanalytic thought, which I have modelled in terms of its derivation from incongruent historical sources. One of these, the Spinozistic, is antithetical to Kantian ways of thinking. This alone does not exclude Longuenesse's proposal, since, I have argued, psychoanalysis is not consistently Spinozistic. It may be true, therefore, that there are aspects of Freud's metapsychology which accord with Kant's concepts. The problem, however, is that psychoanalysis is not consistently Kantian either, and its ambiguity cannot be resolved in either the one direction or the other. This should not, I have urged, be made an objection to psychoanalysis. But if correct, it means that psychoanalysis does not offer a philosophically safe home for Kant's 'I ought' to the extent that Longuenesse supposes.[20]

References

Alexander, F. (1935). 'The logic of emotions and its dynamic background'. *International Journal of Psychoanalysis 16*: 399–413.

Assoun, P.-L. (1976). *Freud: la philosophie et les philosophes*. Paris: Presses Universitaires de France.

Augustine (400/2002). *Confessions*, trans. R. S. Pine-Coffin. Harmondsworth: Penguin.

Carriero, J. (2005). 'Spinoza on Final Causality'. In D. Garber and S. Nadler (eds), *Oxford Studies in Early Modern Philosophy*, Vol. 2 (pp. 105–148). New York: Clarendon Press.

Davidson, D. (2005). 'Spinoza's Causal Theory of the Affects'. In *Truth, Language, and History: Philosophical Essays*, Vol. 5 (pp. 295–314). Oxford: Clarendon Press.

(p. 43) Fichte, J. G. (1794/1982). *The Science of Knowledge, with the First and Second Introductions*, eds and trans. P. Heath and J. Lachs. Cambridge: Cambridge University Press.

Henrich, D. (2003). *Between Kant and Hegel: Lectures on German Idealism*, ed. D. S. Pacini. Cambridge, MA: Harvard University Press.

Kant, I. (1781/1999). *Critique of Pure Reason*, trans. and eds P. Guyer and A. W. Wood. Cambridge: Cambridge University Press.

Kitcher, P. (1992). *Freud's Dream: A Complete Interdisciplinary Science of Mind*. Cambridge, MA: MIT Press.

Koistinen, O. (2009). 'Spinoza on Action'. In O. Koistinen (ed.), *The Cambridge Companion to Spinoza's Ethics* (pp. 167–187). Cambridge: Cambridge University Press.

Lear, J. (1990). *Love and its Place in Nature: A Philosophical Reconstruction of Psychoanalysis*. London: Faber.

Loewald, H. (1988). *Sublimation: Inquiries into Theoretical Psychoanalysis*. New Haven, CT: Yale University Press.

Longuenesse, B. (2012). 'Kant's "I" in "I Ought To" and Freud's Superego'. *Proceedings of the Aristotelian Society* Supplementary Volume *86*: 19–39.

Maimon, S. (1790/2010). *An Essay on Transcendental Philosophy*, ed. N. Midgley, trans. N. Midgley, H. Somers-Hall, A. Welchman, and M. Reglitz. London: Continuum.

Marquard, O. (1987). *Transzendentaler Idealismus, romantische Naturphilosophie, Psychoanalyse*. Cologne: Dinter.

Nicholls, A. and Liebscher, M. (eds) (2010). *Thinking the Unconscious: Nineteenth-Century German Thought*. Cambridge: Cambridge University Press.

Nietzsche, F. (1887/1994). *On the Genealogy of Morals*, trans. C. Diethe, ed. K. Ansell-Pearson. Cambridge: Cambridge University Press.

Novalis (1796/2003). *Fichte Studies*, trans. and ed. J. Kneller. Cambridge: Cambridge University Press.

Redding, P. (1999). *The Logic of Affect*. Ithaca, NY: Cornell University Press.

Ricœur, P. (1977). *Freud and Philosophy: An Essay on Interpretation*, trans. D. Savage. New Haven, CT: Yale University Press.

Sartre, J.-P. (1939/2004). *Sketch for a Theory of the Emotions*, trans. P. Mairet. London: Routledge.

Sartre, J.-P. (1943/1958). *Being and Nothingness: An Essay on Phenomenological Ontology*, trans. H. E. Barnes. London: Methuen.

Schelling, F. W. J. (1800/1978). *System of Transcendental Idealism*, trans. P. Heath, Introduction M. Vater. Charlottesville, VA: University of Virginia Press.

Spinoza, B. (2002). *Complete Works*, trans. S. Shirley, ed. M. L. Morgan. Indianapolis, IN: Hackett.

Tugendhat, E. (1986). *Self-Consciousness and Self-Determination*, trans. P. Stern. Cambridge, MA: MIT Press.

Völmicke, E. (2005). *Das Unbewusste im Deutschen Idealismus*. Würzburg: Königshausen and Neumann.

Wollheim, R. (1972). 'The Mind and the Mind's Image of Itself'. In *On Art and the Mind: Essays and Lectures* (pp. 31–53). Cambridge, MA: Harvard University Press. (p. 44)

Notes:

([1]) For a clear statement of the point, see Tugendhat (1986: 131–2). The puzzle here may be regarded as a heightened, theoretically elaborated form of the perplexity Saint Augustine reports in contemplating the vastness of memory: 'Although it is part of my nature, I cannot understand all that I am. This means, then, that the mind is too narrow to contain itself entirely. But where is that part of it which it does not itself contain? Is it somewhere outside itself and not within it? How, then, can it be part of it, if it is not contained in it? I am lost in wonder when I consider this problem. It bewilders me' (Augustine 400/2002: 216; Book X, 8).

([2]) In what follows I am heavily indebted to Marquard's ground-breaking study (1987), though my reconstruction does not follow his account in all details. Other illuminating studies of the field include Assoun (1976), Redding (1999), and Völmicke (2005). Detailed discussion of major figures in the development of the concept of the unconscious can be found in Nicholls and Liebscher (2010).

([3]) See E3 P13 (Spinoza 2002: 286) and E3 P28 (293).

([4]) See E4 P1 (Spinoza 2002: 323–4), E4 P7 (325–6), and E4 P14 (328–9).

([5]) Thus Davidson (2005) seems to identify causal explanation within the attribute of Thought with rationalization. My construal accords with the interpretation of Spinoza in Carriero (2005: see especially 136–41).

([6]) See E3 P9 S (Spinoza 2002: 284).

([7]) Key points in Spinoza's extended argument against this conception are E2 P48–9 (2002: 272–7), E4 P1–7 (2002: 323–6), and E5 Preface (2002: 363–5).

([8]) A mental decision, Spinoza continues, 'is not distinct from imagination and memory, and is nothing but the affirmation which an idea, in so far as it is an idea, necessarily involves' (E3 P2 S; 2002: 282).

([9]) Which is why achieving freedom is associated with the metaphysical modification of our identity—identification with the immortal part of the mind—described in Part Five.

([10]) Whence Nietzsche's praise of Spinoza for restoring the world's 'innocence' (1887/1994: 50; Essay II, §15). Spinoza describes himself as seeking to 'understand' the emotions and actions of men rather than 'abuse or deride' them (E3 Preface; 2002: 277).

([11]) This is of course very far from a complete analysis, which leaves aside the special case where reason has become determining; my intention is simply to highlight points of relevance to psychoanalysis. On Spinoza's complex theory of agency, see Koistinen (2009).

([12]) See E1 P11, Second Proof (Spinoza 2002: 222). Also relevant is the famous comparison of a person's consciousness of their agency with the thinking of a stone in motion (Letter 58; 2002: 908–10).

([13]) Of interest here is the sketch of a psychoanalytic logic of affect in Alexander (1935), which attempts to coordinate talk of 'emotional syllogisms' with the modelling of mental processes on bodily functions.

([14]) 'But my argument is this: in Nature nothing happens which can be attributed to its defectiveness, for Nature is always the same, and its force and power of acting is everywhere one and the same; that is, the laws and rules of Nature according to which all things happen and change from one form to another are everywhere and always the same' (E3 Preface; Spinoza 2002: 278). Also relevant are E3 P2 S (2002: 281–2), on the cognitive illusion of the distinctness of mental decision, and the deflationary remarks on perfection in the Preface to Part Four (2002: 102–3).

([15]) Kitcher (1992) provides a sophisticated version of this approach.

([16]) As Ricœur puts it, the 'aporia' of psychoanalysis, 'the whole problem of the Freudian epistemology', may be 'centralized in a single question' concerning the integration of 'interpretation that necessarily moves among meanings' with explanation by forces, and 'Freudianism exists only on the basis of its refusal of that disjunction' (Ricœur 1977: 66–7).

([17]) Sartre does not, however, reject symbolic interpretation as such, and towards the end of *Being and Nothingness* he proposes what he considers a superior conceptualization (1943/1958: 602–15).

([18]) Deconstructionist and neo-Nietzschean readings of Freud and applications of psychoanalytic ideas, referred to earlier, construe psychoanalytic theory as engaged in (negative) metaphysical assertion. This removes the problem under discussion, since if at the heart of Freud's theory of the mind *stands the claim* that the concept of the person or subject is incoherent and/or that the reality in which persons and subjects consist does not permit coherent conceptualization, then the theory is freed from the tension I have been exploring and the inconsistency alleged by Sartre: the tension and inconsistency are relocated in the concept which is the theory's object. What this approach may be right about is that psychoanalysis counts among its data the difficulty which subjects

encounter in conceiving themselves in the terms that psychoanalysis recommends, in other words, in relating themselves to their unconscious states, but it is not plausible to construe Freudian theory, as a whole, as a consistently sceptical metatheory.

([19]) Loewald concludes, after consideration of the 'reconciliation of the subject–object dichotomy' involved in sublimation (1988: 20), that psychoanalysis refers us to a *natura naturans* characterized by a subjectivity vaster than 'human individual mentation' (1988: 79–80). Also of high interest is the discussion in Lear (1990: ch. 5).

([20]) Reprinted from Sebastian Gardner (2015), 'Psychoanalytic Theory: A Historical Reconstruction', *Aristotelian Society Supplementary Volume 86*(1): 41–60, Copyright © 2015, Oxford University Press, by permission of Oxford University Press.

Sebastian Gardner

Department of Philosophy, UCL, UK

Oxford Handbooks Online

From Recognition to Intersubjectivity: Hegel and Psychoanalysis 🔒

Molly Macdonald

The Oxford Handbook of Philosophy and Psychoanalysis
Edited by Richard G.T. Gipps and Michael Lacewing

Abstract and Keywords

This chapter engages in a conversation between two seemingly disparate discourses of Hegelian philosophy and psychoanalytic theory. In the work of Jean Hyppolite, the chapter locates the first use of the term intersubjectivity in relation to Hegel's theory of recognition as found in Phenomenology of Spirit (1977) and argues that Hyppolite pioneered a mode of thinking about Hegel and psychoanalysis. The chapter traces the transmutation initiated by Hyppolite through the work of Paul Ricouer, Jacques Lacan, Judith Butler, Jessica Benjamin, and others to locate the influence of Hegel's philosophy on a particular line of psychoanalytic thinking. The author's reading of Hegel incorporated here argues that the majority of attention paid to the concept of recognition, and its translation into terms of intersubjective relating, has confined it to the episode of the Lord and Bondsman and that further attention should be paid to the drama of the Unhappy Consciousness which follows.

Keywords: Hegel, psychoanalysis, recognition, intersubjectivity, Phenomenology of Spirit, unhappy consciousness, Jean Hyppolite, Paul Ricoeur, Jessica Benjamin

Introduction

I would like to illuminate a particular conversation, initiated in the first part of the twentieth century and still alive today, which takes place between the two seemingly disparate discourses of Hegelian philosophy and psychoanalytic theory. This conversation can take a number of different forms, the most recognizable of which involves a focus on the work of Jacques Lacan or on concepts such as negation, as in the work of Slavoj Žižek, following Lacan.[1] My own focus, however, is on exploring the concept of intersubjectivity as it rises out of the Hegelian concept of recognition, which is arguably one of the richest and most significant concepts found in the *Phenomenology of Spirit*. The theory of intersubjectivity was introduced by Jean Hyppolite to describe this process of recognition, to name the relationship between two consciousnesses.[2] Intersubjectivity comes to describe what takes place in the movement from consciousness into self-consciousness, dramatized by Hegel in the Lordship and Bondage episode, as a movement that relies on the interaction, the interrelating of two subjects, each responsible not only for themselves, but also for the other with whom they are bound. The notion of the 'subject', freighted as it is with its own particular history in philosophical and theoretical discourse, asks on the most basic level that, rather than see our existence as anchored in a stable and concrete self, we reconceive of ourselves as subjects, constantly in the process of being made, a product of both our internal lives and our interactions with the world of others and objects. For the purposes of the work I'm doing here to think of Hegel together with psychoanalysis, I would like to consider the Hegelian subject along the lines suggested by Judith Butler (1999: xv). She sees this subject as 'ek-static'—as a 'subject who constantly finds itself outside itself, and whose periodic expropriations do not lead to a return to a former self'. This is a self deeply involved in constant (re)formation. In this process, it is misrecognition as much as recognition that constitutes the subject, and which continuously causes the loss of the self, a 'perpetual displacement' complementing a perpetual desiring and bound subject. In other words, this state of being outside itself is the very thing that drives the Hegelian subject's existence and is never a permanent state, but only part of the dialectical movement. A fundamental part of the dialectical movement of consciousness is that the subject is returned to itself through its interactions with another self-consciousness.

To find a concept of intersubjectivity in Hegel, then, as Hyppolite does and as I do in following him, is not to distort Hegel's original intention, but serves only to recast what is already there: we are, for Hegel, a process of becoming, and this process entails being bound, and unbound, and rebound to another consciousness, as much as we are to the world of objects. Indeed, it is this idea of process and emergence of the self through its relations with the world of objects and others that will serve as one of the key links between Hegel's thought and psychoanalytic theory and practice. As psychoanalyst André Green (1999: 41) writes: 'something is continuing which started with Hegel: the idea of consciousness as a process which is self-revelatory through its discourse, the path from latency to actuality'. One of the very basic tenets, both practical and theoretical, of

Subscriber: Tulane University; date: 17 December 2018

psychoanalysis is the communication (discourse) of the analysand's unconscious, to themselves and to the analyst. This is a process through which the patient brings to self-consciousness (self-reveals) that which has remained hidden, repressed, as yet unthought, but deeply felt and therefore known in some affective form. The crucial aspect of what takes place in an analytic encounter is that this process is initiated, witnessed, interpreted by and with the analyst. We can take this witnessing as a form of recognition and thus contend that the encounter in the analytic setting is a unique form of intersubjective relating that leads to the elaboration of consciousness. Through the lens of Hegel, I read Green here to mean that to move from latency to actuality is to move from simply being a perceiving, sense-certain consciousness, to a dynamic, truly alive, intersubjectively formed self-consciousness. Out of the analytic encounter is where, in the language of Hegel's *Phenomenology*, new shapes of knowing can come to life, manifesting in new moments of self-consciousness.

What follows in this chapter is first a location of the concept of recognition in Hegel's *Phenomenology of Spirit.* I work at extending its reach beyond the Lordship and Bondage episode,[3] through the shapes of consciousness which Hegel dramatizes in the figures of the Stoic and the Sceptic and into a consideration of the Unhappy Consciousness. I offer a reading of the Unhappy Consciousness as a way of remedying what I see as the limitations and blind spots of distilling a concept of recognition, and the notion of intersubjectivity that comes to correspond to it, down to the master/slave dialectic. The figure of the Unhappy Consciousness, too, has its own genealogy in twentieth-century philosophy, starting with the pioneering work of Jean Wahl. Following Wahl, I will argue it is a crucial figure, the 'motor of the dialectic', necessary to break theories of self-consciousness out of the deadlock of master/slave dialectic.[4] I will then turn to the transmutation of the concept of recognition into that of intersubjectivity, touching upon examples of what I see as the most significant interventions in this move. In so doing I cement a positioning of the idea of intersubjectivity as one of the most powerful links to be made between Hegel's philosophical work and contemporary psychoanalytic theory.

Recognition in the *Phenomenology of Spirit*

The Process of Being Bound to the Other: Lordship and Bondage

In the translation of the concept of recognition into that of intersubjectivity performed earlier by Jean Hyppolite we see his equation of the 'need of desire for recognition in another desire' and the term 'intersubjectivity'.[5] I now locate this material in Hegel's *Phenomenology* itself. In Chapter B 'Self-Consciousness', in the section titled 'Independence and Dependence of Self-Consciousness: Lordship and Bondage' Hegel (1977: 111) writes:

Subscriber: Tulane University; date: 17 December 2018

> Self-consciousness exists in and for itself when, and by the fact that, it so exists for another; that is, it exists only in being acknowledged. The Notion of this its unity in its duplication embraces many and varied meanings. […] The detailed exposition of the Notion of this spiritual unity in its duplication will present us with the process of Recognition.

The elaboration of this process of recognition involves both the loss of the self in an other and then the return to the self *via* this other. Here, when self-consciousness 'is faced' by another self-consciousness, it not only finds itself 'as an *other* being' (and we can recall here Butler's formulation of the Hegelian subject as constituted by being outside of itself) but also negates the other independent being to see only itself in this other. This is a process of loss of the self and return to the self, of negation of the other and affirmation of the self in this negation. Importantly, both self-conscious independent beings are experiencing this process simultaneously: 'Each is for the other the middle term through which each mediates itself with itself and unites with itself; each is for itself, and for the other, an immediate being on its own account, which at the same time is such only through this mediation. They *recognize* themselves as *mutually recognizing* one another' (*PS:* 112).

As the section progresses, the drama of the scene heightens and it is in this dramatic tension that most readings of the *Phenomenology* revel, for what transpires is that this mutual recognition cannot sustain itself in the abstract. In order for self-consciousness to truly come to life there must be more at stake, and what is at stake is life itself:

> In so far as it is the action of the *other*, each seeks the death of the other. But in doing so, the second kind of action, action on its own part, is also involved; for the former involves the staking of its own life. Thus, the relation of the two self-conscious individuals is such that they prove themselves and each other through a life-and-death struggle. They must engage in this struggle, for they must raise their certainty of being *for themselves* to truth, both in the case of the other and in their own case. And it is only through the staking of one's life that freedom is won. […] The individual who has not risked his life may well be recognized as a *person*, but he has not attained the truth of this recognition as an independent self-consciousness (*PS*: 113–14).

To truly be alive, not just a body, a person who exists, one must stake one's life in relating to an other. Freedom only comes about through this life-and-death struggle. This is not, at least not yet, an actual staking of life, but a staking of psychic life. But again, Hegel increases the tension. He posits that this is a one-sided and unequal relationship. Reciprocity collapses in an inevitable splitting into the lord (the independent consciousness who lives for itself) and the bondsman (the dependent consciousness who lives for another). We are introduced to the idea that the two consciousnesses then become mediated by an object that is the product of the work of the bondsman, an object that the lord relies on for his existence. This is an 'object of desire' (*PS*: 115) and it belongs to the bondsman. It becomes, therefore, the medium through which he is both

bound to the lord and potentially free. Importantly, the lord is also completely bound to the bondsman through this object and achieves his recognition through this other consciousness. There are many further moves that Hegel makes in this section, most crucially to do with the nature of desire and of the work of the bondsman, through which he becomes 'conscious of what he truly is' (*PS:* 118). It is through this work, through feeling the absolute fear of death and transforming it with its work that the bondsman thus 'acquires a mind of its own' and thus a kind of freedom that is more free than the Lord's. Yet this is a freedom that is 'still enmeshed in servitude' and Hegel needs to bring us through the subsequent shapes of consciousness in order to elucidate the process of recognition in its various permutations.

There is a certain limitation to the readings of intersubjectivity that stay within the Hegelian confines of the 'struggle for recognition' as based in this moment of the dialectic, even though there are compelling reasons for most interpretations of Hegel's thinking to entrench themselves here and, indeed, the confluence of his thought with psychoanalytic theory is given grounding in this dramatic passage. Part of the responsibility for the lasting impact of this section of Hegel's *Phenomenology* is down to the way in which it was read by Alexandre Kojève whose Marx-inflected lectures on Hegel were to prove remarkably influential on a cohort of French students, who would themselves go on to be leading intellectuals (amongst others George Bataille, Maurice Blanchot, Maurice Merleau-Ponty, Andre Breton, and Lacan). From Kojève on there was a particular flavour of Hegel that was accepted, the Hegel of desire and the 'struggle to the death'. Judith Butler (1997: 11) argues that, because it was perfectly suited to a 'liberationist narrative for various progressive political visions[,] most readers have neglected to pay attention to the resolution of freedom into self-enslavement at the end of the chapter'.[6] I will now briefly move through the stages of the Stoic and the Sceptic before landing at the unhappy consciousness, the site of this self-enslavement. I believe that Butler does not go far enough to see that there is a resolution to this self-enslavement embedded in the episode of the unhappy consciousness, a resolution that brings us back to a scene of intersubjective relating.

Stoicism and Scepticism

The state of externalization and freedom reached at the end of the Lordship and Bondage section quickly transforms into the internalized state of *thought* of the Stoic as self-consciousness attempts to understand its freedom. Consciousness begins to lose sight of the necessity of the other and the 'I' becomes a seemingly self-sufficient being: 'In thinking, I *am free*, because I am not in an *other*, but remain simply and solely in communion with myself, and the object, which is for me the *essential* being, is in undivided unity being-for-myself; and my activity in conceptual thinking is a movement within myself' (*PS*: 120). The process of recognition here is solipsistic, self-consciousness involved only in recognizing itself. This recognition then changes further into the shape of Scepticism as thought is actualized. From the externalization of consciousness which results in a bounded state, both to its own object and to the external world, what looks like a move towards freedom is a consciousness edging slowly towards the unbounded state. In scepticism the freedom of stoicism 'becomes a reality', 'duplicates itself, and now knows itself to be a duality'. We have here a remarkable splitting of the self into 'dual natured, merely contradictory being' (*PS*: 126). This splitting causes deep suffering for self-consciousness and this deeply suffering state is the unhappy consciousness.

The Unhappy Consciousness

Building on the work of Wahl, Hyppolite (1974: 190) proposed that 'Unhappy consciousness is the fundamental theme of the *Phenomenology*. Consciousness, as such, is in principle always unhappy consciousness, for it has not yet reached the concrete identity of certainty and truth, and therefore it aims at something beyond itself'. Such a reading does not follow the traditional 'end of history' readings of the *Phenomenology* but instead argues that even when one gets to the 'end', one is always still aiming at 'something beyond'. This is the movement of consciousness, of the dialectic itself.

The birth of the unhappy consciousness is, in many ways, the manifestation, the actualization, of the potential for the pain and suffering that are the potential for all who possess self-consciousness. It is a product of the movements of the stoic and the sceptic that have come before. It is the culmination of the fear and the terror of the 'life-and-death struggle' that took place with the lord and bondsman. It is, in some senses, the affective response of consciousness articulated and made apparent since we have now turned to the idea of the activity of consciousness as it occurs in one individual rather than between two. Here we encounter consciousness that is experiencing a crisis of existence. Unlike the sceptic that set about freeing itself, with a miscalculated coldness, from itself and from the world in which it existed, the attitude of the unhappy consciousness is not one of indifference. If anything, it could be accused of feeling too intensely its own existence.

What the unhappy consciousness is bound to at this stage is the *feeling* of its own meaninglessness. Hegel writes: 'Consciousness of life, of its existence and activity, is only an agonizing over this existence and activity, for therein it is conscious that its essence is only its opposite, is conscious only of its own nothingness' (*PS*: 127). The essence of its activity is the *desire* for its own self as this essential being. What it finds is that what it seeks, the essential, the 'other' in an individual form, will necessarily always be out of its reach. Ultimately, the lack it feels, the *desire* to fill this lack, is its activity. We can return to Butler's formulation of the Hegelian subject as that which is always somehow outside itself, a kind of becoming subject spurred on by its loss of self. Hegel writes of this despair in order, I think, to alert us to the danger of unbinding ourselves from the world of others, to show us the deep vulnerability that is the price of believing that freedom means complete and utter independence. But he also asks us to think about the vulnerability of our desires.

In a remarkable move in this section, Hegel introduces the figure of a 'counsellor', the third, to save the Unhappy Consciousness from itself, reinforcing the necessity of an 'other' in order to ground the self, to ground the movement of self-consciousness and to save this newly liberated consciousness from self-despair (*PS*: 136–8). Intersubjectivity, then, is cemented as a mode of being that is fundamental to the survival of self-consciousness, and its realization does not sit with the struggle between master and slave, but in the acknowledgement of this necessity as it rises from a moment of existential crisis. I suggest that when interpreting the *Phenomenology* in this way, if we perform a secular reading of this section, we can see the figure of a proto-analyst (counsellor) emerge, which can bolster a reading of Hegel as anticipating the psychoanalytic project in an unexpected way. In its moment of crisis, consciousness gives itself over to this mediating figure and in so doing returns to itself by relinquishing what it thought was its freedom, but what was really just its frenzy of self-conceit. Once this figure of 'the third' as counsellor is introduced, self-consciousness is saved from itself and returns to Reason. It is not difficult to read an analytic scene into this episode. The crucial move is not that the counsellor has more power than the despairing self-consciousness, but that it is equipped to help the unhappy consciousness see itself in a new light, to form a new mode of knowing, to help relieve the misery and suffering, and to help it continue on in this process of becoming.

From Recognition to Intersubjectivity—On the Way to the Analytic Setting

In order to begin a transition into the analytic setting, I now turn to Paul Ricoeur's masterful work, *Freud and Philosophy* (1970). Ricoeur is not only useful in helping to further nuance the translation performed by Hyppolite of Hegel's process of recognition into the language of intersubjectivity, but he also explicitly locates the analytic encounter as an intersubjective relation. Celebrated, amongst other things, for its coining of the phrase 'the masters of suspicion', Ricoeur's text is one of the most exhilarating works of philosophy written about psychoanalysis, ranging into and beyond an interrogation of the potential resonance between Freudian psychoanalytic theory, hermeneutics, and phenomenology. His reading of Hegel and intersubjectivity which sits embedded in its vast and ranging arguments calls for a reconsideration.[7] While it is not on the surface devoted entirely to Hegel's philosophy, Ricoeur's engagement with Hegel's *Phenomenology* configures it as essential to a productive illumination of Freud's thought.[8] Ricoeur (1970: 468) wishes 'to show in each discipline of thought, considered in and for itself, the presence of its other'. One of his major criticisms of Freud's topographical model (as tied in various ways to the realm of the drives/instincts), which he accuses of unhelpful abstraction, is that it 'does not account for the intersubjective nature of the dramas forming its main theme'. He goes on to argue that:

> Whether it be the drama of the parental relation or the drama of the therapeutic relation itself, in which the other situations achieve speech, what nourishes analysis is always a debate between consciousnesses. […] Stated bluntly, the Freudian systematization is solipsistic, whereas the situations and relations analysis speaks of and which speak in analysis are intersubjective

> (Ricoeur 1970: 61).

Ricoeur is here positioning the therapeutic relation as one through which all other 'dramas', the 'situations' (e.g. of family relations, the life of the drives, the desires and wishes etc. of the patient) can finally find articulation and interpretation, finding life in language, in the speech of the patient. This speech is received, contained, and translated back to them by the analyst, hopefully effecting the relief from suffering that brought them into analysis in the first place. This is the potential power of the intersubjective relating which takes place in an analysis.

Ricoeur is clearly indebted here to the Lacanian location of speech as the foundational act of intersubjectivity, and the move beyond the solipsism of the drive is necessary in order to privilege speech as the centre of the analytic encounter. In his early seminars, the concept of intersubjectivity seems central to Lacan's work and his indebtedness to Hegel is acknowledged, particularly in his first Seminar (1953).[9] Lacan argues that 'Speech is the founding medium of the intersubjective relation', and locates hate as one

pole (as the opposite of love) of the 'very structure of the intersubjective relation'. He remarks that this is 'what Hegel recognises as the impasse for the coexistence of two consciousnesses, from whence he deduces his myth of the struggle for pure prestige' (Lacan 1991: 274–7). Language, then, specifically speech, is that which mediates intersubjective relating, both within and outside of the analytic setting.

Lacan's contribution to thinking about intersubjectivity here centres on questions of language and his contribution is to translate the Hegelian process of recognition into a process more fitting for the analytic setting, in that it becomes a form of recognizing the speech of the other. Here we can return to the scene of the Lord and Bondsman to suggest that what Lacan is picking up on is the idea that for Hegel the relationship between two self-consciousnesses is mediated by the 'object of desire', the 'thing' that then holds them both in bondage, bound to one another and to the thing itself. It is worth noting that in the early chapters of the *Phenomenology* Hegel puzzles through the role of language in the experience of consciousness and most importantly the failure of language to adequately express one's encounters with the world (the experience of a sensing and perceiving consciousness), yet he does not really engage with the question of language when he comes to think through the movement of self-consciousness.[10] As we've seen, he is concerned with desire, its motivation and manifestation, with the fashioning of an object out of the self, of the activity of thought (recall the Stoic). This brings us to acknowledge the crucial idea that it is largely left up to Hegel's reader to render the drama of recognition into terms that draw it out of abstraction and into the material world. So, for example, there is plenty of opening for Marx to recast the scenes of the lord and bondsman as the drama of the worker and the owner of capital. The mediating object becomes then the product of the worker's labour. For Lacan, we can see his investment in structuralist linguistics and in language translate the mediating objects to that of Speech.

Yet, here we might want to pause just for a moment and consider the manner in which the nature of non-verbal recognition, the fact that the analytic encounter begins with two bodies in a room (one potentially prone on a couch and the other sitting behind them, a strange and vulnerable physical asymmetry) is also a basis for intersubjective relating. Simultaneously, where, we might want to wonder, are the bodies in Hegel's *Phenomenology*?[11] It takes some time to think through the body with Hegel, but it is there in the idea of work and desire, and also in an admission of the 'animal functions' (*PS*: 135). We should not lose sight of the body, even as we encounter a surfeit of affect (despair, fear, 'the struggle of the heart and emotions', the 'feeling of wretchedness' (*PS*: 135)). How bodies are the site of recognition is worth considering.

To return to Ricoeur then with the idea of transference: as could be expected, this becomes the locus of intersubjective relating in the analytic setting, whose possibility 'resides in the intersubjective texture of desire and of the desires deciphered within that situation' (Ricoeur 1970: 372). It is important to note Ricoeur's contribution here, as it is, as far as I can tell, completely neglected by those 'schools' of relational and intersubjective psychoanalysis which have come to define a strand of the contemporary

American scene (to whom I will turn my attention shortly), even though the corrective they offer to the Freudian corpus is the very same Ricoeur offered in 1970. His reading of intersubjectivity comes as part of the inheritance of Hyppolite's Hegelian line of thinking, and he discusses this at some length, with particular reference to Hyppolite's paper 'Phenomenologie de Hegel et Psychanalyse' mentioned earlier, moving through the 'texture' of desire, and the formation of subjectivity out of the intersubjective, the analytic relationship (Ricoeur 1970: 386–9). In fact, Ricoeur goes so far as to say that there is a structural homology between Hegel's dialectic and the process of consciousness as it arises in the analytic setting: 'The entire analytic relation can be reinterpreted as a dialectic of consciousness, rising from life to self-consciousness, from the satisfaction of desire to the recognition of the other consciousness' (Ricoeur 1970: 474) Ricoeur does not, however, explicitly move away from using the master/slave episode of the dialectic. In a slightly troubling manner, he at first traps the analytic situation in the unequal relationship of this dynamic, positioning the patient in the role of the slave, whose truth lies in the other, 'before becoming the master through a work comparable to the work of a slave, the work of the analysis'. Yet, he also offers that the way out of this impasse is that the patient/slave becomes the master through the difficult and demanding labour of analysis. The analyst as master is simply a stage in the dialectic of analysis. The logical outcome of this work for Ricoeur is that the analysis will end when there is equality between the two consciousnesses when 'the truth of the analyst has become the truth of the sick consciousness' and is no longer alienated, and has become a self, himself. (Ricoeur 1970: 474) The implication of this reading is not only to securely position the analytic encounter as intersubjective, but also to suggest that the analyst, although seemingly in the position of mastery, is just as vulnerable as the patient. We can see how Ricoeur is recasting the Hegelian drama as an analysis. As we saw with Hegel, the lord (master) is as bound to the bondsman (slave) for their existence as the latter to the former. We also might want to see in Ricoeur's interpretation an analysis which culminates, without explicitly being stated here, in the move from the self-divided alienation of the unhappy unconsciousness, to the form of Reason.

There is one final element of Ricoeur's deeply nuanced reading to attend to and that is the strong note of caution he offers. He acknowledges that although it is through the concept of intersubjectivity that we can see the closest identification between phenomenology and psychoanalysis, it is also the site of their most radical distinction. The fundamental difference is that psychoanalysis treats the 'intersubjective relationship as *technique*' (Ricoeur 1970: 406, italics his). On a basic level (and he does spend some time working through what technique means for psychoanalysis—from the technique of treatment, methods of investigation, and the body of theories at play) Ricoeur breaks the word 'technique' down into thinking about the 'work' of an analysis as having three parts: the analyst's work, the work of the analysand ,and then, the most slippery of all, the mechanisms of the neurosis. The intersubjective relation of a psychoanalysis—the analytic dialogue—because it is a technique that relies on the technique of transference interpretation, is unlike any other form of intersubjectivity. To forget this, or gloss over

this, is to do a disservice to both psychoanalysis and intersubjectivity in their fullest potentials and different registers.

Hegel and Intersubjectivity in Contemporary Psychoanalysis

Arnold H. Modell (1993) doubts 'whether our present psychology of intersubjectivity could have developed without Hegel. For Hegel appeared to have intuitively grasped the fundamental aspects of the psychology of self and other'. Indeed, Modell continues, 'Hegel can justifiably be termed the first intersubjective or relational psychologist' (Modell 1993: 98–9). In the last decades of the twentieth century there was a movement towards intersubjective theory on the side of psychoanalytic theory written from a more clinical perspective. Given the scope of this chapter, I cannot fully address the complexities of the contested nature of such developments in psychoanalytic theory or trace the nuances of the way in which the American strands of the psychoanalytic institutional sphere differ from those of the UK, France etc.[12] I wish here only to pick up some of the threads of my discussion of the translation of Hegel's process of recognition into the theory of intersubjectivity and to see where this translation might find resonance in the field of contemporary psychoanalysis.

Henry Schwartz (2012) has argued that the introduction of intersubjective theory has been seen as a declaration of war by the more traditional American school of ego psychology, yet that what the various modes of theorizing and practice that could come under the umbrella term of intersubjectivity (relational, interpersonal, intersubjective systems theory), are trying to effect is a corrective to the Freudian legacy of the focus on the drive and intrapsychic life.[13] I have already noted earlier that Ricoeur set this corrective in motion in his own work, but clearly Ricoeur was not a practising psychoanalyst, and it is not entirely surprising that the American branches of the psychoanalytic project would not engage with his thinking. Following André Green's lead, Schwarz notes that psychoanalysis has always been intersubjective but that the contemporary movements are doing the work of elaborating upon that which has existed from the start without yet being properly theorized. That is to say, the corrective in post-Freudian thinking has been to bring together theories of the drive with theories of the object, to think about intersubjective relating, and not just intrapsychic life, and to crucially consider the dialectic at play between the drive and the object.[14]

Although it is not the aim of this chapter to explore fully the nuances of the battles taking place in intersubjective theory, it is important for an understanding of where Hegel is still at play to note the basic distinctions. The predominately American 'intersubjective' and 'relational' modes of analysis had their origins in 1970s American psychoanalytic movements. Yet, it is only the relational school that could be considered as having institutional boundaries (with, as Schwarz notes, a doctoral training programme at New

York University started in 1988, and a professional journal) whereas we can consider the intersubjective to be a mode of theorizing that finds itself in many strands of psychoanalytic practice. The relational school is mainly associated with the work of Stephen Mitchell and Jay Greenberg and of the two schools is of less interest to my survey because it is not invested in a Hegelian model of intersubjectivity. It had its origins in Mitchell and Greenberg's work *Object Relations in Psychoanalytic Theory* (1983). The best-known representatives of one development of the intersubjective school are Robert Stolorow, George Atwood, and Donna Orange. Atwood and Stolorow claim their first use of the term 'intersubjectivity' was in an article in 1978, but develop theories of intersubjectivity throughout their works such as Atwood and Stolorow's *Structures of Subjectivity: Explorations in Psychoanalytic Phenomenology* (1984), *Working Intersubjectively: Contextualism in Psychoanalytic Practice* (1997), written with Orange, and *Worlds of Experience: Interweaving Philosophical and Clinical Dimensions in Psychoanalysis* (2002).

These authors define intersubjectivity as a 'sensibility that continually takes into account the inescapable interplay of the two subjects in any psychoanalysis. It radically rejects the notion that psychoanalysis is something one isolated mind does to another, or that development is something one person does or does not do' (Orange, Atwood, and Stolorow 2002: 18). Although they take Hegel's master/slave model as the starting point for any kind of thinking about 'dominance relations' and for any thinking about the relation of self and wider historical and intellectual contexts, they do not use the master/slave paradigm to represent a developmental stage of subjectivity, as does Jessica Benjamin, a psychoanalyst who explicitly pays her debt to Hegel's model, but claims to draw her use of the term 'intersubjectivity' out of Habermas's work, rather than, as I have, Hyppolite's.[15]

Benjamin's work marks a second important strain of intersubjectivity theory that uses Hegel and the master/slave paradigm as a model for recognition and intersubjective relations and for what is seen as grounding an understanding of the dialectical nature of the theory and practice of psychoanalysis. Three of her works are key here: *The Bonds of Love: Psychoanalysis, Feminism and the Problem of Domination* (1990), *Shadow of the Other: Intersubjectivity and Gender in Psychoanalysis* (1998) and, 'Beyond Doer and Done To: An Intersubjective View of Thirdness' (2004). Benjamin shows a consistent allegiance to Hegel, but seems to confine herself to his most famous chapter of the *Phenomenology*.

As a way of situating an element of Benjamin's thinking around intersubjectivity let me turn to the idea noted earlier of mutual recognition as part of a developmental process. We are here witnessing the translation of Hegel's process of recognition into terms about human psychic development. Benjamin is interested in the notion of the infant's development of intersubjectivity as it derives from the recognition of its carers (e.g. the mother). She is fully aware of that critique of her work which posits that she 'conflates the normative ideal of mutual recognition with the empirical possibility and necessity of development'.[16] Yet, she argues it is not that the baby begins with a normative ideal, but

'with a material possibility', a view which has been excluded from the Freudian view of the baby and thus of all human development as seen in strictly Freudian theory. She writes:

> Thus, mutual recognition derives from material possibility, but whether it will be realized and in what way are in no sense predetermined. The idea is not to bolster the ideal by proving that we are 'born with the ability'; it is to recognize that when we postulate a psychological *need* (not a social need or a normative ideal) for recognition, we mean that failure to satisfy the need will inevitably result in difficulties or even damage to the psyche

> (Benjamin 1995: 20).[17]

We can understand Benjamin here in light of our earlier consideration of Hegel's *Phenomenology* if we think about Hegel's argument that one can be in existence, but that without recognition, one is not truly living. In Benjamin's terms this comes to be about a healthy and thriving psyche being reliant upon the process of being recognized. Benjamin is equally aware, however, of the critique of a collapsing of the tension between the dialogue between the psychoanalytic and the philosophical register when talking recognition and (inter)subjectivity. In *Like Subjects, Love Objects* she rehearses an argument made in her earlier work *The Bonds of Love*:

> On the one hand, I stated that we need mutual recognition in order to develop our faculties, a statement that postulates a kind of essential, development position that is typical of psychoanalysis; on the other hand, I argued from the standpoint of social and political theory that we need mutual recognition in order to live in some degree free of domination and non-violently. I linked these two positions by arguing that the upshot of failures of recognition is domination, that the constitution of subjectivity and the self-other relationship is a necessary material basis for noncoercive intersubjectivity

> (Benjamin 1995: 21).

Importantly, for Benjamin, the idea of mutual recognition includes autonomy, which is a point that I think Donna Orange seems to miss in her critique of Benjamin's work noted earlier. Indeed, I think Orange misreads both Hegel and Benjamin. On the one hand, when she gets stuck on the notion of mastery in the process of intersubjective recognition, Orange completely misses the idea of the freedom of the bondsman through work and the creation of the object and the resolution of the drama of master and slave into the later stages of self-consciousness. So, this fundamental misreading of Hegel also leads Orange to misread Benjamin's intentions in thinking with Hegel about domination, mastery, recognition, and the 'mastery' of the analyst.[18] Benjamin is thoroughly Hegelian in her vision of mutual recognition in which there is a necessary tension between dependence and independence for the subject and in the requirement for the inclusion of the possibility for breakdown, of a failure to sustain the tension as well as to account for the possibility of repair after this failure. Benjamin does not do so, but we could read this

as an engagement with the movement of self-consciousness out of the scene of the master and slave and into the other stages I rehearsed above of the stoic, sceptic, and the unhappy consciousness.

The final point I want to make regarding Benjamin's work is to emphasize not just the theoretical aspect of the concept of (mutual) recognition, but also the practical, where she draws it into the analytic scene. For Benjamin (1995: 24), it is productive to think about intersubjective theory not just as a supplement or corrective to intrapsychic theory, but to acknowledge it as a 'concrete site of discovery and exploration', a site that 'has become central to thinking about the analytic endeavour', especially around the notion of counter-transference and the particular subjectivity of the analyst as it comes to meet with the mind of the patient. Here are echoes of the argument I traced through Ricoeur and of my own suggestions about how we can form a dialogue between Hegel's philosophy as it is found in the *Phenomenology* and psychoanalytic theory. It is, at its heart, a dialogue that offers an understanding of what it means to be a subject involved in the continuous process of becoming, as bound to another subject as part of this process—whether inside or outside the unique intersubjective encounter that is an analysis. To be recognized and to recognize in return is one of most difficult but necessary elements of being human.

References

Agamben, G. (1991). *Language and Death: The Place of Negativity*, trans. K. E. Pinkus and M. Hardt. Minneapolis, MN: University of Minnesota Press.

Atwood, G. E. and Stolorow, R. D. (1984). *Structures of Subjectivity: Explorations in Psychoanalytic Phenomenology*. Hillsdale, NJ: The Analytic Press.

Benjamin, J. (1990). *The Bonds of Love: Psychoanalysis, Feminism and the Problem of Domination*. London: Virago.

Benjamin, J. (1995). *Like Subjects, Love Objects*. New Haven, CT: Yale University Press.

Benjamin, J. (1998). *Shadow of the Other: Intersubjectivity and Gender in Psychoanalysis*. London: Routledge.

Benjamin, J. (2004). 'Beyond Doer and Done To: An Intersubjective View of Thirdness'. *Psychoanalytic Quarterly 73*: 5–46.

Benjamin, J. (2010). 'Can we Recognize each other? Response to Donna Orange'. *International Journal of Psychoanalytic Self Psychology 5*(3): 244–256.

Butler, C. (1976). 'Hegel and Freud: A Comparison'. *Philosophy and Phenomenological Research 36*(4): 506–522.

Butler, J. (1997). *The Psychic Life of Power*. Stanford, CA: Stanford University Press.

Butler, J. (1999). *Subjects of Desire: Hegelian Reflections in Twentieth-Century France*, New York: Columbia University Press. (Original work published 1987.)

Christensen, D. (1968). 'Hegel's Phenomenological Analysis and Freud's Psychoanalysis'. *International Philosophical Quarterly 8*(3): 356–378.

Derrida , J. (1978). *Writing and Difference*, trans. A. Bass. Chicago, IL: Chicago University Press.

Derrida, J. (1982). *Margins of Philosophy*, trans. A. Bass. Brighton: Harvester Press.

Dews, P. (2007). *Logics of Disintegration*. London: Verso.

Green, A. (1999). *The Work of the Negative*. London: Free Association Books.

Green A. (2000). 'The intrapsychic and intersubjective in psychoanalysis'. *Psychoanalytic Quarterly 69*: 1–39.

Habermas, J. (1993). 'Labour and interaction: Remarks on Hegel's Jena *Philosophy of Mind*'. In *G. W. F. Hegel: Critical Assessments,* 4 vols. London*:* Routledge. (Original work published 1967.)

Hegel, G. W. F. (1977). *Phenomenology of Spirit*, trans. A. V. Miller. Oxford: Oxford University Press. (Original work published 1807.)

Hyppolite, J. (1969). 'The Human Situation in the Hegelian Phenomenology'. In *Studies in Marx and Hegel*, trans. J. O'Neill. London: Heinemann. (Original work published in 1955.)

Hyppolite, J. (1971). 'Hegel's Phenomenology and Psychoanalysis'. In W. Steinkraus (ed.), *New Studies in Hegel's Philosophy*, trans. A. Richer. New York: Holt, Rinehart & Winston. (Original work published in 1957.)

Hyppolite, J. (1974). *Genesis and Structure of Hegel's Phenomenology.* Chicago, IL: Northwestern University Press. (Original work published in 1946.)

Hyppolite, J. (1997). *Logic and Existence*, trans. L. Lawlor and A. Sen. Albany, NY: SUNY Press. (Original work published in 1952.)

Kojève, A. (1980). *Introduction to the Reading of Hegel*, assembled by R. Queneau, ed. A. Bloom, trans. J. Nichols. Ithaca, NY: Cornell University Press.

Lacan, J. (1988). *The Seminar of Jacques Lacan: Book I, Freud's Papers on Technique 1953–1954*, ed. J.-A. Miller, trans. J. Forrester. Cambridge: Cambridge University Press

Lacan, J. (1991). *The Seminar of Jacques Lacan: Book II, The Ego in Freud's Theory and in the Technique of Psychoanalysis 1954–1955*, ed. J.-A. Miller, trans. S. Tomaselli. New York: Norton.

Macdonald, M. (2014). *Hegel and Psychoanalysis: A New Interpretation of the Phenomenology of Spirit*. London: Routledge.

McCumber, J. (1993). *The Company of Words: Hegel, Language and Systematic Philosophy*. Evanston, IL: Northwestern University Press.

Mead, G. H. (1962). *Mind, Self and Society*, ed. C. Morris. Chicago, IL: University of Chicago Press. (Original work published 1934.)

Mitchell, S. A. and Greenberg, J. R. (1984). *Object Relations in Psychoanalytic Theory*. Cambridge, MA: Harvard University Press.

Modell, A. (1993). *The Private Self*. Cambridge, MA: Harvard University Pres.

Ogden, T. (1994). *Subjects of Analysis*. Northvale, NJ: Jason Aronson.

O'Neill, J. (ed.) (1996). *Hegel's Dialectic of Desire and Recognition*. Albany, NY: SUNY Press.

Orange, D. (2010). 'Recognition as: Intersubjective Vulnerability in the Psychoanalytic Dialogue'. *International Journal of Psychoanalytic Self Psychology 5*(3): 227–243.

Orange, D. M., Atwood G. E., and Stolorow, R. D. (1997). *Working Intersubjectively: Contextualism in Psychoanalytic Practice*. Hillsdale, NJ: The Analytic Press.

Orange, D. M., Atwood G. E., and Stolorow, R. D. (2002). *Worlds of Experience: Interweaving Philosophical and Clinical Dimensions in Psychoanalysis*. New York: Basic Books.

Reid, J. (2007) *Real Words: Language and System in Hegel*. Toronto/London: University of Toronto Press.

Ricoeur, P. (1970). *Freud and Philosophy: An Essay on Interpretation*, trans. D. Savage. New Haven, CT: Yale University Press. (Original work published in 1965.)

Roudinesco, E. (1990). *Jacques Lacan & Co. A History of Psychoanalysis in France 1925–1985*, trans. J. Melman. Chicago, IL: University of Chicago Press.

Roudinesco, E. (1999). *Jacques Lacan: An Outline of a Life and a History of a System of Thought*, trans. B. Bray. Cambridge: Polity Press.

Russon, J. (1997). *The Self and its Body in Hegel's Phenomenology of Spirit*. Toronto: University of Toronto Press.

Schwartz, Henry P. (2012). 'Intersubjectivity and dialecticism'. *International Journal of Psychoanalysis 93*: 401–425.

Surber, J. O. (ed.) (2006). *Hegel and Language*. Albany, NY: SUNY Press.

Veron, J. (2007). *Hegel's Philosophy of Language*. London: Continuum.

Wahl, J. (1993). 'The Unhappy Consciousness in the Philosophy of Hegel'. In *Hegel: Critical Assessments*, trans. R. Northey (pp. 284–310). London: Routledge. (Original work published in 1951.)

Williams, R. (1992). *Recognition: Fichte, Hegel and the Other*. Albany, NY: SUNY Press.

Williams, R. (1997). *Hegel's Ethics of Recognition*. Berkeley, CA: University of California Press.

Žižek, S. (1993). *Tarrying with the Negative: Kant, Hegel And The Critique Of Ideology*. Durham, NC: Duke University Press.

Žižek, S. (1996). *The Indivisible Remainder: An Essay on Schelling and Related Matter*. London: Verso.

Žižek, S. (2005). *Interrogating the Real*, eds R. Butler and S. Stephens. London: Continuum.

Žižek, S. (2012). *Less Than Nothing: Hegel and the Shadow of Dialectical Materialism*. London: Verso.

Notes:

(1.) I will not engage with Žižek's work here, but see Žižek (1993), (1996), and a collection of his essays (2005) (where he posits the 'negation of negation', as the 'Hegelian version' of 'death drive' (34)) has offered, in no systematic fashion, a set of readings that mobilizes Hegel in order to understand Lacan and vice versa. He consistently argues that Lacan himself was thoroughly Hegelian, even after he renounced his philosophy, and even, as Žižek argues, if he didn't know it himself. His work, *Less Than Nothing: Hegel and the Shadow of Dialectical Materialism* (2012) offers many brilliant readings of Hegel. See Part III, Section 8 'Lacan as a Reader of Hegel' (pp. 507–51); section 11 'Negation of the negation: Lacan Versus Hegel?' (pp. 787–94). Although Lacan's relationship with Hegel's philosophy is mentioned in many texts, for the most succinct account from the point of view of an intellectual biography see Elisabeth Roudinesco, *Jacques Lacan & Co. A History of Psychoanalysis in France 1925–1985* (1990) and Elisabeth Roudinesco, *Jacques Lacan: An Outline of a Life and a History of a System of Thought* (1999), pp. 61–107.

(2.) Hyppolite is perhaps best known in his role as translator of the *Phenomenology of Spirit* into French (1939–41) and for his highly influential *Genesis and Structure of Hegel's Phenomenology* (1946, 1979), and *Logic and Existence* (1952, 1997). However, Hyppolite contributed two ground-breaking pieces on the subject of Hegel, psychoanalysis and the concept of intersubjectivity. The first is a lecture given in 1955 to the French Society for Psychoanalysis titled 'Hegel's Phenomenology and Psychoanalysis',

which was transcribed by Jean Laplanche and first published in French as 'Phenomenologie de Hegel et Psychanalyse' in *La Psychanalyse, Vol III* (1957). English language readers were given access to it in a collection on Hegel titled *New Studies in Hegel's Philosophy*, edited by Warren E. Steinkraus (1971). The version of the paper included in the Steinkraus volume was given reconsideration by Hyppolite himself, including the addition of illuminating footnotes as to his discussions with Lacan following the delivery of the original lecture on the relationship between Freud and Hegel, and the implications for the crossover of the disciplines of psychoanalysis and philosophy. The second is 'The Human Situation in the Hegelian Phenomenology' in *Studies in Marx and Hegel* (1955, 1969).

(3.) The Lordship and Bondage episode is often referred to as the dialectic of master and slave by the majority of twentieth- and twenty-first-century philosophical texts, as a result of following Alexandre Kojève's influential reading, which I discuss later. See *Introduction to the Reading of Hegel* (1980). When reading the Hegel, I will refer to it in his language of Lordship and Bondage, but when engaging with other readers of Hegel, I will refer to it as they do.

(4.) Wahl's seminal text has yet to be published in English, bar a short excerpt from *The Unhappy Consciousness in the Philosophy of Hegel* (1993) which has curtailed his influence.

(5.) Nearly all of the scholarly texts on the *Phenomenology* discuss the concept of recognition. It is important to note, however, that, as I've suggested, very few translate it into terms of intersubjectivity. Although he does not acknowledge Hyppolite's original use of the term intersubjectivity in regards to Hegel, Robert Williams's *Recognition: Fichte, Hegel and the Other* (1992) also claims that recognition (*Annerkennung*) is the ground of intersubjectivity, but that the concept of intersubjectivity 'in German idealism is all but unnoticed' (1992: 1). Williams credits Fichte and Hegel with having brought the concept to philosophy and spends his work detailing the inception and progression of the concept of intersubjectivity. His aim is to recover this original ground for what he locates as a burgeoning interest in the philosophy of intersubjectivity and the 'problem of the Other' taking place in twentieth-century intellectual thought. See also his work *Hegel's Ethics of Recognition* (1997) which is a continuation of this earlier work. For a collection of articles that discuss Hegel along similar lines, see John O'Neill (ed.), *Hegel's Dialectic of Desire and Recognition* (1996).

(6.) The influence Kojève's lectures on Hegel of 1933–39 had on thinkers like Lacan and his contemporaries has been well documented. See Roudinesco, *Lacan* & Co., pp. 134–47; Peter Dews, *Logics of Disintegration* (2007), pp. 64–6; p. 304, n26. Robert Williams argues in *Hegel's Ethics of Recognition* (1997) that Kojève's narrow focus, which equates recognition with the struggle between master and slave, 'distorted' Hegel's original concept of intersubjectivity.

(7.) In addition to Ricoeur's work, there are two articles of note that were published in roughly the same era as Ricoeur's 1970 publication, and are important for understanding the way in which the confluence of Hegel and psychoanalytic theory started to elaborate itself. Neither deal directly with the concept of intersubjectivity so will not be addressed here. See Darrell Christensen's, 'Hegel's Phenomenological Analysis and Freud's Psychoanalysis', in the *International Philosophical Quarterly* (1968) and Clark Butler's 'Hegel and Freud: A Comparison' in *Philosophy and Phenomenological Research* (1976). One can see in Butler's article an indication of just how undeveloped the line of inquiry into the Hegel/psychoanalysis connection is at that point of the twentieth century with a glance at the bibliography. There are six titles, one being Christensen's article, four of which had yet to be translated into English, including the article by Hyppolite, 'Phenomenologie de Hegel et la Psychanalyse'. When Butler writes in the first line of his article that, 'Exploration has recently begun into the relation between Freud and Hegel' he is sitting at a particular juncture of history where this kind of thinking would have started to be acceptable outside of Lacanian circles.

(8.) [Eds: See also Bernstein, this volume].

(9.) The majority of Lacan's commentary on Hegel can be found in *Seminar I,* particularly Chapter XVII 'The object relation and the intersubjective relation'. Of additional interest is *Seminar of Jacques Lacan*: *Book II*, *The Ego in Freud's Theory and in the Technique of Psychoanalysis 1954–1955*, Chapter XV, 'Odd or even? Beyond intersubjectivity'.

(10.) For interesting engagements with the question of Hegel and language see Jacques Derrida, 'The Pit and the Pyramid: Introduction to Hegel's Semiology', in *Margins of Philosophy* (1982), pp. 69–108, and 'Force and Signification' in *Writing and Difference* (1978) pp. 3–30. Georgio Agamben offers an interesting meditation in *Language and Death: The Place of Negativity* (1991). Also see Jeffrey Reid, *Real Words: Language and System in Hegel* (2007); Jim Veron, *Hegel's Philosophy of Language* (2007); *Hegel and Language* (2006); John McCumber, *The Company of* (1993).

(11.) For an important study of this question see John Russon, *The Self and its Body in Hegel's Phenomenology of Spirit* (1997). Russon also follows Hyppolite in locating the Unhappy Consciousness as the heart of the *Phenomenology.*

(12.) This is as distinct from what might be considered object-relations theory, deriving from the work of Donald Winnicott and geographically located more on the UK side of the Atlantic. There is great potential to read Hegel alongside other branches of psychoanalytic thinking that come after Winnicott, and I do so in the portion of my book *Hegel and Psychoanalysis: A New Interpretation of the Phenomenology of Spirit* (2014) concerning the work of Christopher Bollas.

(13.) [Eds: For further discussion of the debate between traditional and intersubjective schools on psychoanalysis, see also Lacewing this volume.]

(14.) See also Green (2000).

(15.) Habermas's 'Labour and interaction: Remarks on Hegel's Jena *Philosophy of Mind*' (1967) draws a major distinction between Fichte and Hegel regarding the foundation of the 'I'. For Fichte, Habermas argues, the foundation of the 'I' comes in the idea of the 'subjectivity of self-knowing' and this creates the dialectic between the 'I' and the 'other, whereas for Hegel, the framework is that of 'intersubjectivity of spirit, in which the "I" communicates not with itself as its "other" but instead with another "I" as its "other"' (Habermas 1993: 560). This is a key moment in the trajectory of the concept of intersubjectivity, because from this point on Habermas is often credited with bringing it into twentieth-century theory, both via G. H. Mead's (1934) reading of mutual recognition and directly from his own interpretation of Hegel, whereas Hyppolite is somewhat overlooked as a point of origin. Interestingly, however, Habermas does not simply fix his reading of mutual recognition and thus of intersubjectivity in the master/slave dialectic in the 1807 *Phenomenology* as most versions of the theory do, but instead finds the roots in Hegel's Jena lectures of 1803–4 and 1805–6.

(16.) See *International Journal of Psychoanalytic Self Psychology*, Vol 5, Number 3, 2010 for an in-depth and discussion and engagement between Donna Orange, Jessica Benjamin, and others around the issue of intersubjectivity as drawn from theories of (Hegelian) recognition in the clinical setting. This special issue is dedicated to an attempt to find common ground between the self psychologists, 'intersubjective system theorists' (noted as Stolorow, Atwood, Orange) and interpersonal and relational analysts. It notes the historical dissonance about the concept of recognition within the schools of psychoanalytic theory and practice listed here and stems from a critique launched by Orange in her 2008 article 'Recognition as: Intersubjective Vulnerability in the Psychoanalytic Dialogue' (reprinted in this issue) that Benjamin's Hegelian model of recognition puts the analyst in the position of mastery and does not allow for a necessary vulnerability on their part.

(17.) Thomas Ogden's work (1994) is similarly representative of this line, and owes and acknowledges a great debt both to Hegel's *Phenomenology* and to André Green's work. The majority of the theory found in Ogden's *Subjects of Analysis* is shot through with the model of dialectical tension (subject and object, analyst and analysand) that he sees as inherent in subjectivity.

(18.) Benjamin (2010) herself responds to Orange's (2010) misreadings in a collegial yet forceful rejoinder in her article 'Can we Recognize each other? Response to Donna Orange'.

Molly Macdonald

School of English and Drama, Queen Mary University of London

Oxford Handbooks Online

Abstract and Keywords

This chapter engages in a conversation between two seemingly disparate discourses of
Hegelian philosophy and psychoanalytic theory. In the work of Jean Hyppolite, the
chapter locates the first use of the term intersubjectivity in relation to Hegel's theory of
recognition as found in Phenomenology of Spirit (1977) and argues that Hyppolite
pioneered a mode of thinking about Hegel and psychoanalysis. The chapter traces the
transmutation initiated by Hyppolite through the work of Paul Ricouer, Jacques Lacan,
Judith Butler, Jessica Benjamin, and others to locate the influence of Hegel's philosophy
on a particular line of psychoanalytic thinking. The author's reading of Hegel
incorporated here argues that the majority of attention paid to the concept of recognition,
and its translation into terms of intersubjective relating, has confined it to the episode of
the Lord and Bondsman and that further attention should be paid to the drama of the
Unhappy Consciousness which follows.

Keywords: Hegel, psychoanalysis, recognition, intersubjectivity, Phenomenology of Spirit, unhappy consciousness,
Jean Hyppolite, Paul Ricoeur, Jessica Benjamin

Andrew Brook and Christopher Young

Introduction

THE will and the unconscious were among the most important subjects in the German-
speaking intellectual world in the nineteenth century. This line of inquiry may well have
reached its highest development later in Freud but many of the doctrines developed by
Freud did not begin with him, or even with Nietzsche. To find their first clear expression
we must go back at least as far as Arthur Schopenhauer. In Schopenhauer, we find not
only ideas similar to some of Freud's most characteristic ideas but a surprisingly
complete articulation of them. It is generally known that some of Freud's views are

similar to some of Schopenhauer's.[1] Indeed, later in life Freud himself acknowledged this, though with an ambivalence to which we will return. However, the correspondences are far more extensive and detailed than have been previously recognized, at least in English. We will explore these correspondences[2] and examine Freud's attitude to Schopenhauer.

Will and Sexuality

We turn first to Schopenhauer's concept of the will. Though primarily a metaphysical concept, it is also at the root of Schopenhauer's explicitly psychological doctrines, (p. 64) especially his view that sexuality is pervasive in all human motivation and that intellect is secondary to the will. For Schopenhauer, the will is fundamental. It underlies and animates everything that we can observe. According to Schopenhauer, individual volition is a limited (also indirect and distorted) manifestation of the same will from which the entire observable world arises. We can know something of will's manifestation in us but we cannot be conscious of it directly. In Schopenhauer's pessimistic view, the will's endless strivings are forever incapable of satisfaction. Setting aside the broader metaphysical functions that Schopenhauer assigns to the will, let us examine how he saw its manifestations in the volition of individual human beings.

Schopenhauer thought that the will itself is a blind unconscious striving which is not directly accessible to our own consciousness. However, we enjoy a partial awareness of it in our own volition, and it further manifests itself in sexual desire and in our 'love of life'. These are all manifestations of an underlying will to live. Freud's first theory of instinct, a picture in which dual instincts are rooted in a single will to live, is very similar and it remained his view until his *Beyond the Pleasure Principle* (1920), i.e. for the first two decades of his psychoanalytic writings. For both thinkers, the sex drive was by far the stronger of the two, 'the most perfect manifestation of the will to live' (Schopenhauer 1844 (2): 514).[3] Indeed, Schopenhauer went so far as to claim that:

> man is concrete sexual drive; for his origin is an act of copulation, and his desire of desires is an act of copulation, and this impulse alone perpetuates and holds together the whole of his phenomenal existence. (1844 (2): 514)

> The sexual impulse is the most vehement of all craving, the desire of desires, the concentration of all our willing.[4] (1844 (2): 514)

> To all this corresponds the important role which the sex-relation plays in the world of mankind, where it is really the invisible central point of all action and conduct, and peeps up everywhere in spite of all the veils thrown over it. It is the cause of war and the aim and object of peace, the basis of the serious and the aim of the joke, the inexhaustible source of wit, the key to all allusions, and the meaning of all mysterious hints, of all unspoken offers and all stolen glances; it is the daily meditation of the young and often the old as well, the hourly thought of the unchaste, and even against their will the constantly recurring imagination of the

chaste, the ever ready material for a joke, just because the profoundest seriousness lies at its root. (1844 (2): 513, translation slightly modified)

(p. 65) On how people respond to the dominating force of sexual desire, Schopenhauer again sounds like Freud. His account of how far human beings will go to deny the power of sexuality is every bit as acerbic as Freud's:

> This … is the piquant element and the jest of the world, that the chief concern of all men is pursued secretly and ostensibly ignored as much as possible. But, in fact, at every moment we see it seat itself as the real and hereditary lord of the world, out of the fullness of its own strength, on the ancestral throne, and looking down from thence with scornful glances, laugh at the preparations which have been made to subdue it, to imprison it, or at least to limit it and if possible to keep it concealed, or indeed so to master it that it shall only appear as a subordinate, secondary concern of life.

> (Schopenhauer 1844 (2): 513)

The passages just quoted are so completely in line with psychoanalysis that it is easy to forget that their author was dead by the time Freud entered school as a young boy. Even Freud did not give sexual desire any stronger a role in human motivation than Schopenhauer did.

The parallels go deeper. They both greatly expand the range of things to which the concept of sexuality applies, and in much the same way. Both stretched sexuality far beyond procreation and many of Schopenhauer's illustrations of the disguised manifestations of sexual desire could have been written by Freud. They both come close to using the term to describe virtually all bodily pleasure-seeking of any sort, though Freud went further than Schopenhauer in this regard.

Although Schopenhauer drastically broadened the range of motive and activity called 'sexual', to the point where not much sexual in the ordinary sense could be found, unlike Freud he always kept some link to the orgasmic, the genital—to sexuality in its ordinary sense. If the will is the ground of everything, including all instincts and therefore something much broader than normal sexuality, at least for him its *manifestations* are sexual in the ordinary sense. Freud went further: he not only expanded the range of the sexual, he expanded the concept itself, calling many things sexual that have no obvious link to the orgasmic or genitalia at all.[5]

Finally, both Schopenhauer and Freud treated sexuality from two very different perspectives: the individual and the species (but see Sandford, this volume). As Schopenhauer wrote, 'It is true that the will-to-live manifests itself primarily as an effort to maintain the individual; yet this is only a stage towards the effort to maintain the species' (1844 (2): 514). Here is the same dual perspective in Freud: 'On the one view, the individual is the principal thing, sexuality is one of its activities and sexual satisfaction is one of its needs; while on the other view, the individual is a temporary and

transient appendage to the quasi-immortal germ-plasm, which is entrusted to him by the process of generation' (1915a: 125).

(p. 66) Freud did depart most notably from Schopenhauer in his views about infantile sexuality. Contrary to what Freud later held, Schopenhauer attributes the happiness of youth to the fact that the sexual impulse, so 'pregnant with evil, is lacking in the child … ; from this arises the character of innocence, intelligence, and reasonableness' which we find in children (1844 (2): 395).[6] This may be because Schopenhauer linked sexual desire to reproduction more closely than Freud did (Sandford, this volume).

In light of this disagreement about infantile sexuality, it is interesting that the two of them did agree on the central importance of childhood to adult life. As Freud put the point, 'the child is psychologically father to the adult and … the events of his first years are of paramount importance for his whole later life' (1940a: 187). There are few ideas for which Freud is better known. It is not well known that Schopenhauer may have held the same view:

> the experiences and acquaintances of childhood and early youth become thereafter the types and rubrics of all later knowledge and experience … Thus the firm foundation of our world view is formed even in the years of childhood, together with its shallowness or depth: it is later carried out and completed; yet not essentially altered.
>
> (quoted in McGill 1971: 55)

Returning to the will and instincts, Freud parallels Schopenhauer on some more theoretical issues. They were both unsettled as to whether there is one fundamental kind of motivator or two, an unease that came to a head for Freud in 1920. Schopenhauer often distinguished the sexual drive from 'the love of life' as two drives ('*Next* to the love of life', he once said, sexual love is 'the strongest and most active of all motives … ' (1844 (2): 533, our emphasis)). But he also often ran them together in an undifferentiated notion of the will. That he saw them, even when separated, as both manifestations of will also has a parallel in Freud: Freud saw both libido and the self-preservative drive as discharge of endogenous stimuli or avoidance of excessive exogenous stimuli.

Will or Drives and Intellect

We turn now to the relation of the will to the intellect. According to Schopenhauer, the will must objectify itself in the world to satisfy its strivings. To do so, it creates for itself an intellect appropriate to its needs. Even memory relies on the unifying force of the will's stable and unceasing urges: (p. 67)

> if we reflect deeply on the matter, we shall reach the conclusion that memory … requires the foundation of a will as … a thread on which the recollections range themselves, and which holds them firmly together, or that the will is, so to speak,

the ground on which the individual recollections stick, and without which they could not be fixed.

(Schopenhauer 1844 (2): 222)

We will have more to say about the thread of memory in the sections on 'Psychopathology' and 'Psychic Defences'. Here the important point is that even this element of the intellect is secondary to the will and subordinate to its demands. Being the basis of the intellect, the will 'rules it, guides it, incites it to further effort, in short imparts to it the activity that is not originally inherent in it' (Schopenhauer 1844 (2): 213, 224). The will creates the intellect, which 'is designed merely to prescribe to the individual will its motivations, i.e. to indicate to it the objectives of its desires together with the means of taking possession of them' (Schopenhauer 1970: 59). This led Schopenhauer to the idea that the intellect was not as rational as had been previously supposed; the will dictates, unseen, what the mind desires, believes, and thinks. Our states of consciousness and our decisions had previously been thought to be the outcome of processes of reasoning. Schopenhauer argues that these states have their origin in the will.

Schopenhauer was not the first to part company with the Enlightenment view that reason dominates, but he gave the will far greater prominence than any previous thinker, building his whole model of the psyche upon it. Freud, of course, shared Schopenhauer's view that the 'intellect is entirely secondary' to the functioning of the mind: *'the ego is not master in its own house'* (Freud 1917: 143, italics in original). Further, they both saw that the intellect promptly takes the demands of the will as its own: 'the ego is in the habit of transforming the id's will into action as if it were its own' (Freud 1923: 25). There are even shadows of the concept of rationalization in Schopenhauer. He did not explicitly formulate the concept but, like Freud, he held the view that the intellect takes what are really the will's motives as its own and justifies them as though it were the author of them (Schopenhauer 1970: 59). Here is Freud's version of roughly the same view:

We are bound to suppose that a unity comparable to the ego cannot exist … from the start; the ego has to be developed. (1914b: 76)

We may well … conclude that instincts and not external stimuli are the true motive forces behind the advances that have led the nervous system, with its unlimited capacities, to its present high level of development. (1915a: 120)

For Freud, activities of what he called the ego are in the service of an (un)pleasure principle (1895), in service of discharging energy that is causing unpleasure, and he eventually developed a sophisticated account in which discharges are not outwards but at the self, an activity he called auto-eroticism and needed in order to make room for the concept of narcissism (1914b: 88). So his account goes beyond Schopenhauer's. Nevertheless, the two views start from the same picture.

Schopenhauer and Freud

Freud adhered to this picture in one form or another throughout his life, from the *Project for a Scientific Psychology* (1895) and Chapter 7 of the *Interpretation of Dreams* (1900: 565ff., 598ff.) at least to *Beyond the Pleasure Principle* (1920: 10ff.) and even later, (p. 68) with appropriate modifications to accommodate the introduction of the death instinct. As we have seen, the parallels with Schopenhauer are close. They even agree that the 'afflux of stimulation' that gets mental life going is 'incessant and unavoidable' (Freud 1915a: 120) and that the only way to stop it is to find some object that quiets its source, something that creates an 'experience of satisfaction', for example food or sexual discharge. Schopenhauer's characterizations of how the will operates even anticipate the notion of primary process.

> For what the bridle and the bit are to an unmanageable horse, the intellect is to the will in man; it must be led by this bridle by means of instruction, exhortation, training, and so on; for in itself the will is as wild and impetuous an impulse as is the force appearing in the plunging waterfall; in fact it is, as we know, ultimately identical therewith. (1844 (2): 213)

Parallels even extend to their respective views of pleasure and the way the will operates. Both saw pleasure as merely negative, a removal of an irritant. Pleasure is, for Schopenhauer the momentary cessation of the will's striving, for Freud the discharge or at the very least the achievement of constancy in the flow of stimuli from the drives. 'Every unpleasure' ought 'to coincide with a heightening, and every pleasure with a lowering, of mental tension due to stimulus' (Freud 1924a: 159–60). Only in 1924 did Freud even partially modify this view. Thus, for the first twenty-five years of his work on psychology, he adhered to the view of Schopenhauer, whether or not he was aware of it. Since obviously much pleasure is not like this, it is interesting that both argue (or assume) that it is, though others have, too.

Schopenhauer may have carried through the implications of the primacy of the will or the id for rationality more consistently than Freud. If Freud was a child of the German Romanticism to which Schopenhauer so richly contributed, he was also a child of nineteenth-century scientific empiricism. In line with the latter, he believed that the human mind could operate rationally and discover truths about the world. If rationality was threatened by the unconscious, it was a threat that could be overcome, at least in science. From this point of view, when Schopenhauer said that '[e]very passion, in fact every inclination or disinclination, tinges the objects of knowledge with its colour … most common of occurrence is the falsification of knowledge brought about by desire or hope' (1844 (2): 141), he was perhaps more in tune with this implication of the power of the will or id than Freud was. As some psychoanalysts in France and elsewhere demonstrate, Freud's model of the mind can easily be taken in this direction, too.

Unconsciousness

Let us dwell for a moment on the specific idea that we are conscious only of expressions of the will in individual volition, not the will itself. The parallels with Freud on this issue are particularly close. Says Schopenhauer, the intellect 'does not penetrate into the secret workshop of the will's decisions'. Indeed: (p. 69)

> the intellect remains so much excluded from the real resolutions and secret decisions of its own will that sometimes it can only get to know them, like those of a stranger, by spying out and taking unawares; and it must surprise the will in the act of expressing itself, in order merely to discover its real intentions.

> (Schopenhauer 1844 (2): 209–10)

Compare Freud:

> Unconsciousness is a regular and inevitable phase in the processes constituting our psychical activity; every psychical act begins as an unconscious one, and it may either remain so or go on developing into consciousness, according as it meets with resistance or not.

> (Freud 1912: 264)

These parallels in the two doctrines of the unconscious may not seem especially surprising to us now, standing as we do on the other side of Freud. A further parallel is striking even now: Schopenhauer's main argument for the existence of unconscious mental states is also the argument that Freud used most frequently. The argument begins by noting that much thought, feeling, and behaviour cannot be explained on the basis of conscious mental states alone. 'It is evident', wrote Schopenhauer, 'that human consciousness and thinking are by their nature necessarily fragmentary' (1844 (2): 138). By this he means that conscious psychological states often seem disjointed, do not add up to any coherent picture of what the beliefs, feelings, and motives behind a belief or action were like. Freud makes exactly the same point:

> [C]onscious acts remain disconnected and unintelligible if we insist upon claiming that every mental act that occurs in us must also necessarily be experienced by us through consciousness. (1915c: 167)

Yet both took for granted that there is a unity, a coherence, to mental life. What produces it? Schopenhauer: 'What gives unity ... to consciousness is ... the will. The will ... holds all its ideas and representations together ... ' (1844 (2): 139–40). Compare Freud:

> We have found—that is, we have been obliged to assume—that very powerful mental processes or ideas exist ... which can produce all the effects in mental life that ordinary ideas do ... though they themselves do not become conscious. (1923: 14)

Freud used this argument repeatedly throughout his career (1909: 175–6; 1915c: 166ff.; 1923: 14–18; 1940: 196–7). It is very similar to Schopenhauer's argument.

Relation to Consciousness

The two thinkers also agree on important aspects of the relationship of the unconscious to consciousness. Schopenhauer believed just as strongly as Freud that vastly more of the psyche is unconscious than conscious. 'Consciousness is the mere surface of our (p. 70) mind, and of this, as of the globe, we do not know the interior, but only the crust' (1844 (2): 136). Again:

> Let us compare our consciousness to a sheet of water of some depth. Then the distinctly conscious ideas are merely the surface; on the other hand, the mass of the water is the disposition of our … will that is the kernel of our inner nature. (1844 (2): 135)

Or, as he put it in a work that Freud cites three times in the *Interpretation of Dreams*, 'the intellect is a mere superficial force, essentially and everywhere touching only the outer shell, never the inner core of things' (1851: 301). The same image of consciousness as an outer crust or cap runs through the whole of Freud's work, from the *Project* to the famous diagram of the mind in Lecture 31 of the *New Introductory Lectures* (1933: 78). The parallel is made more striking by the fact that they both held that only this outer layer of the psyche is in contact with or affected by the external world. The intellect, 'enlightened by experience … draws up and modifies its precepts', says Schopenhauer (1844 (2): 224), but the will is unaffected by external reality. In Freud, the 'processes in the id are entirely unconscious, while consciousness is the function of the ego's outermost layer, which is concerned with the perception of the external world' (1925b: 266).

Freud even shared Schopenhauer's notion that consciousness is not the natural state of psychological contents, that bringing these contents to consciousness is an extra step. Schopenhauer: 'It is no more possible for an idea to enter consciousness without an occasion than it is for a body to be set in motion without a cause' (1844 (2): 133). Freud:

> every psychical act begins as an unconscious one, and it may either remain so or go on developing into consciousness, according as it meets with resistance or not (1912: 264, a passage also cited earlier).[7]

For Schopenhauer, what makes it possible to follow these paths of association is that ideas, i.e. psychological states, are arranged in an ordered sequence, along temporal, causal, and narrative lines, tied together in virtue of being successive stages in the unfolding of the unchanging projects of the will. This continuity is also what makes possible the thread of memory introduced in the section on 'Will or Drive and Intellect'. What associations do, to simplify a bit, is to allow us to trace these sequences in various ways.

Schopenhauer and Freud

Schopenhauer's views on association are also similar to the associative mechanisms that Freud identified as central to dream work and psychic defence, namely, condensation and displacement (Freud 1900: Ch. 4$_A$ to 4$_D$). Schopenhauer even uses the example of remembering a forgotten dream to illustrate his theory.

> The search for the thread of recollection shows itself in a peculiar way, when it is a dream that we have forgotten on waking up. Here we look in vain for that which a (p. 71) few minutes previously occupied us with the force of the clearest and brightest present, but has now entirely vanished. We then try to seize any impression that has been left behind, and on which a slender thread hangs. By virtue of association, this thread might draw the dream back into our consciousness. (1844 (2): 134)

This passage expresses the core of Freud's later notion of free association. Freud of course broadened both the theoretical formulation and the applications of the idea, but the basic notion is already there in Schopenhauer.

To be sure, their views of consciousness and the unconscious may seem to differ in some respects. Schopenhauer emphasizes his view that the will is 'the true and ultimate point of unity of consciousness, and the bond of all its functions and acts' (1844 (2): 140). By unity, he seems to have meant the continuity and stability of our basic drives and interests. Freud would have denied that what Schopenhauer called the will has even that much unity, being governed as it is by the primary process. However, the disagreement here may be more apparent than real. In Schopenhauer, when the will is manifested in individual persons, its unity is lost. In this state, Schopenhauer, too, sees it as pursuing conflicting aims without regard to mutual contradiction—a cardinal feature of the primary process.

A real difference is that Schopenhauer always viewed consciousness, the intellect, as a separate psychic system. Freud went both ways on the question. He could say that consciousness is 'only a *quality* or attribute of what is psychical, and moreover an inconsistent one' (1940b: 285–6). He could also tie it to the preconscious and treat the two together as a separate psychic structure, the structure that with further development evolves into the ego.

To summarize, there are striking parallels between Schopenhauer and Freud both in their doctrines of the will or id and in its relationship to consciousness. For both, the will or the id is unconscious and consists of non-rational processes, seeks satisfaction endlessly and insatiably, provides the most powerful motives in human life, mostly unknown, causes rationality and the conscious mind to come to be, and is needed to explain thought, feeling, and action. For both, these views force a fundamental re-evaluation of consciousness.

Psychopathology

We turn now to a quite different topic—Schopenhauer and psychopathology. Schopenhauer thought that madness is caused by the repression of painful memories or traumata, though he did not use the word 'repression' (see the next section entitled 'Psychic Defences'). By 'madness', he seems to have meant what we would now call psychosis or severe affective disorder. As an account of psychosis, his theory is not very plausible. However, as an account of neurosis, it is very similar to Freud's first theory of neurosis: 'hysterics [neurotics of one kind] suffer mainly from reminiscences' (1893–5: 7). (p. 72) Schopenhauer introduces his observations with the penetrating remark that madness does not affect the entire range of mental capabilities:

> Neither the faculty of reason nor understanding can be denied to the mad, for they talk and understand, and often draw very accurate conclusions. They also, as a rule, perceive quite correctly what is present, and see the connection between cause and effect.

> (1819 (1): 192)

Thus, these disorders lie in something other than a simple inability to connect with reality. In fact, they originate in problems of memory:

> For the most part, mad people do not generally err in knowledge of what is immediately present; but their mad talk relates always to what is absent and past, and only through these to its connection with what is present. Therefore, it seems to me that their malady specifically concerns the memory (1819 (1): 192).

We are reminded of Freud's first theory. But there is more. For Schopenhauer, the failure of memory that causes the trouble is something very specific. In some cases memory is preserved, but in others:

> it is a case of the thread of memory being broken, its continuous connection being abolished, and of the impossibility of a uniformly coherent recollection of the past. Individual scenes of the past stand out correctly, just like the individual present; but there are gaps in their recollections that they fill up with fictions … In his memory the true is forever mixed up with the false. Although the immediate present is correctly known, it is falsified through a fictitious connection with an imaginary past.

> (1819 (1): 192)

As we recall, Schopenhauer used the metaphor of the thread of memory to describe the ordering of consciousness by the will (see the sections entitled 'Will or Drives and Intellect' and 'Psychic Defences', in this chapter); when the thread is broken, there are periods in the will's activities that are no longer represented in memory, the continuity of aim in these activities notwithstanding. Given what Schopenhauer said about repression, he probably held, as did Freud later, that while they are lost to consciousness, they are

Subscriber: Stockholm University; date: 16 March 2019

not lost altogether. What could break the thread? According to Schopenhauer in a remarkable passage, traumata are the villains.

> The fact that violent mental suffering or unexpected and terrible events are frequently the cause of madness, I explain as follows. Every such suffering is as an actual event always confined to the present; hence it is only transitory, and to that extent is never excessively heavy. It becomes insufferably great only in so far as it is a lasting pain, but as such it is again only a thought, and therefore resides in memory. Now if such a sorrow, such a painful knowledge of reflection, is so harrowing that it becomes positively unbearable, and the individual would succumb to it, then (p. 73) nature, alarmed in this way, seizes upon madness as the last means of saving life. The mind, tormented so greatly, destroys, as it were, the thread of memory, fills up the gaps with fictions, and thus seeks refuge in madness from the mental suffering that exceeds its strength. (1819 (1): 193)

This is all very similar to Freud's first theory noted earlier and to one of Freud's accounts of (at least some kinds) of psychosis: delusions, Freud once said, are 'applied like a patch over the place where originally a rent had appeared in the ego's relation to the external world' (1924b: 151). To be sure, Schopenhauer and Freud were hardly the only ones to believe that hysteria is linked to memory; in the passage in which Freud's famous aphorism occurs, he immediately goes on in a long footnote to discuss others 'who have held similar views on hysteria to ours' (1893–5: 7n). He mentions Moebius, Strümpell, and Benedikt—but not Schopenhauer. Yet Schopenhauer could well have been, directly or indirectly, the source of the idea for all of them.

Schopenhauer was also well aware of the 'psychopathology of everyday life' (Freud 1901) and the connection between mental illness and 'normal' psychological processes. One of Freud's greatest achievements was to make credible the idea that neurotic symptoms are purposive, that there are reasons for them. They are not expressions of arbitrary neurological failure. Once again, we find Schopenhauer arriving at an idea first. Nature, we saw him say, 'seizes upon madness as a means of saving life'. Both Schopenhauer's and Freud's contemporaries mostly saw mental illness as a symptom of neurological breakdown, of constitutional weakness. In sharp contrast, they both saw it as psychologically functional: madness is a way of coping, a motivated reaction. The idea first entered Freud's work about 1890 and he (rightly) never gave it up.

Psychic Defences

In connection with 'madness', Schopenhauer articulated large parts of what became the Freudian doctrine of how the psyche defends itself against painful memories and the like. In particular, though he did not use the word, his insights into the nature of what Freud called repression run very deep. The will always:

makes its supremacy felt in the last resort. This it does by prohibiting the intellect from having certain representations, by absolutely preventing certain trains of thought from arising … it then curbs and restrains the intellect, and forces it to turn to other things. (1844 (2): 208)

Even as late as 1919, Freud ascribed to the same motive for repression as Schopenhauer had: 'we have a perfect right to describe repression, which lies at the basis of every neurosis, as a reaction to a trauma—as an elementary traumatic neurosis' (1919: 210).

(p. 74) Schopenhauer also described what is translated in both thinkers as 'resistance'.[8] We can understand the memory gaps of the mad better, he tells us, when:

we remember how reluctant we are to think things that powerfully prejudice our interests, wound our pride, or interfere with our wishes; with what difficulty we decide to lay such things before our own intellect for accurate and serious investigation; how, on the contrary, pleasant affairs come into our minds entirely of their own accord, and, if driven away, always creep on us once more, so that we dwell on them for hours. *In this resistance on the part of the will to allow what is contrary to it to come under the examination of the intellect is to be found the place where madness can break in on the mind*. (1844 (2): 400, our italics)

Schopenhauer even articulated a thought that is on the road to the clinical insight that it is therapeutic to bring the unconscious to consciousness:

Every new adverse event must be assimilated by the intellect, in other words, must receive a place in the system of truths connected with our will and its interests … As soon as this is done, it pains us much less; but this operation itself is often very painful, and in most cases takes place only slowly and with reluctance. But soundness of mind can continue only in so far as this operation has been correctly carried out each time. (1844 (2): 400)

It is a short step indeed from this claim to the idea of making the repressed unconscious conscious, and from there to the clinical insight that recovering unconscious memories, fantasies, etc. will deprive them of their power by assimilating them into consciousness. In Freud's words, 'The therapeutic influence of psycho-analysis depends on the replacement of unconscious mental acts by conscious ones and is effective within the limits of that factor' (1925b: 265).

As the passages quoted in this and the previous sections show, Schopenhauer had the essentials of the whole traumatic theory of neurosis. He specified the role played by traumatically induced repression of memories, identified the phenomena of substitute fantasies and symptoms, and described resistance fifty years before Freud did.

Other Parallels in Psychological Doctrine

There are other parallels between Schopenhauer's work and Freud's which we do not have the space to explore. Briefly, Schopenhauer had a detailed theory of dreams, a theory that occupies a substantial portion of his hundred-page *Essay on Spirit Seeing and* (p. 75) *Everything Connected Therewith* (1851). Freud cites this essay three times in the *Interpretation of Dreams*. In addition, this essay and some of Schopenhauer's other works display a knowledge of neurophysiology quite remarkable for someone of Schopenhauer's time and discipline. The same was true of Freud, of course, but Schopenhauer's theory of the systems that make up the brain and of how experiences could consist of communication of impulses in and among these systems is actually similar to Freud's view in *The Project* of the psy/phi/omega systems and the Q_n said by Freud to circulate among them. Schopenhauer also described what Freud called parapraxes and sublimation.

Freud on Schopenhauer

The substantial parallels between Schopenhauer and Freud raise a number of questions. The most obvious is the extent to which Freud is *indebted* to Schopenhauer for key insights, whether directly, by having read him, or indirectly, by being exposed to his insights in other ways. For intellectual history is filled with coincidences and not every correlation is the result of a causal connection. We argue that Freud almost certainly was indebted to Schopenhauer. Both Freud's remarks and his silences are of interest here. There is also a further question as to whether Schopenhauer's high standing at the time helped to ease the *reception* of psychoanalysis, and we suggest that indeed it did.

Freud's remarks on Schopenhauer fall into two periods. Before 1915, Freud made few references to Schopenhauer: five in the *Interpretation of Dreams* (Freud 1900) including the three just mentioned, one to a comment about Schopenhauer (Freud 1901: 119), a passage which also contains a reference to von Hartmann, to whom we will return, a footnote on two kinds of knowledge which uses Schopenhauer as an example but could have used anyone (Freud 1909: 196n), and a strange passage in Freud (1914a: 15). Only two of these references are to *The World as Will and Representation*, a general reference to Schopenhauer on time in the *Interpretation of Dreams* (Freud 1900: 503), and the reference in 1914a, which is a passage that Freud knew about at all, he tells us, only because Rank showed it to him. He is arguing in the passage that he discovered repression by himself.

> The theory of repression quite certainly came to me independently of any other source; I know of no outside impression which might have suggested it to me and for a long time I imagined it to be entirely original, until Otto Rank … showed us a passage in Schopenhauer's *World as Will and Representation* in which the philosopher seeks to give an explanation of insanity. [He is probably referring to the passage from 1844 (2): 192–208, quoted earlier.] What he says there about the

struggle against accepting a distressing piece of reality coincides with my concept of repression so completely that once again I owe the chance of making a discovery to my not being well read. (1914a: 15)

(p. 76) Of course, an idea that someone else has already hit upon is not a discovery. And Freud had already read some Schopenhauer, as the references in *Interpretation of Dreams* show. Moreover, the *Essay on Spirit Seeing* (1851) that Freud cites repeats Schopenhauer's memory theory of madness, and his views on dreams, the will, and sexuality, asserts that the intellect is just the shell of the mind, and so on. However, prior to 1915, Freud said little to indicate that he had read any other works by Schopenhauer.

Beginning in 1915, the pattern changes markedly. The *Introductory Lectures* explicitly acknowledge Schopenhauer as a forerunner of psychoanalysis:

> Probably very few people can have realized the momentous significance for science and life of the recognition of unconscious mental processes. It was not psycho-analysis, however, let us hasten to add, which took this first step. There are famous philosophers who may be cited as forerunners—above all the great thinker Schopenhauer, whose unconscious 'will' is equivalent to the mental instincts of psycho-analysis. It was this same thinker, moreover, who in words of unforgettable impressiveness admonished mankind of the importance, still so greatly-underestimated by it, of its sexual craving.

> (Freud 1916–17: 143) [Strachey thinks that Freud had in mind the passage that we cited earlier from 1844 (2): 513–14.]

Then in the 1920 preface to *Three Essays*, we find:

> it has been some time since Arthur Schopenhauer, the philosopher, showed mankind the extent to which their activities are determined by sexual impulses—in the ordinary sense of the word. (1905: 134, 1920 edn)

Freud acknowledged parallels between his thought and Schopenhauer's very fully after 1915.

Yet what he credited Schopenhauer for is a little strange. He credited Schopenhauer for anticipating not his pre-1920 theory of libido and self-preservation as separate drives but his post-1920 theory which merged the two under the concept of Eros and contrasted both with the newly introduced death drive. Indeed, Freud links both parts of his new picture to Schopenhauer. He now treats sexual drives as fundamental to Eros (' … the true life instincts' (1920: 40)), and then links this expanded concept to Schopenhauer (1920: Preface). Similarly, when he introduces his controversial death drive later in 1920, Freud tells us that:

> We have unwittingly steered our course into the harbour of Schopenhauer's philosophy. For him death is the 'true result and to that extent the purpose of life', while the sexual instinct is the embodiment of the will to live. (1920: 49–50)

This is doubly curious. Schopenhauer never married sexuality and the urge to self-preservation in the way that Freud is now doing. And he never postulated a positive drive to die. It was bad enough for him that death was the inevitable *result* of living. In short, (p. 77) Freud first acknowledged the parallels between his drive theory and Schopenhauer's only at the point at which they had largely ceased to exist!

While he now credited Schopenhauer, he continued to assert his originality:

> The large extent to which psycho-analysis coincides with the philosophy of Schopenhauer—not only did he assert the dominance of the emotions and the supreme importance of sexuality but he was even aware of the mechanism of repression—is not to be traced to my acquaintance with his teaching. I read Schopenhauer very late in life. (1925a: 59)

Even if Freud is right, the remark is an evasion. He ignores all the other ways in which he could have learned about Schopenhauer, as many commentators have noted. For example, Herzog noted that 'Freud was fully aware of a philosophical tradition, centred in Germany, that placed great emphasis on the concept of a psychical unconscious . . . [d]espite his insistence that he had read neither Schopenhauer nor Nietzsche until after making his own discoveries' (Herzog 1988: 169). And McGrath (1986: 148) argues convincingly that even if Freud read Schopenhauer only late in life, the parallels on repression might still not be a coincidence. One of Freud's most influential teachers, Meynert, described repression fully and he explicitly credited the idea to Schopenhauer (1851). Freud knew this. Yet as little as a year after Freud wrote the impressive 1914 acknowledgement quoted earlier, he could go so far as to claim that repression was a 'concept which could not have been formulated before the time of psycho-analytic studies' (1915b: 146).

One of the most significant indirect routes by which Freud encountered Schopenhauer's thought is in the works of Eduard von Hartmann, whom we know Freud read early in his career (Brandell 1979: 93). Von Hartmann was a popularizer of Schopenhauer and his *The Philosophy of the Unconscious* (1869) and numerous works of the 1870s were immensely popular. Schopenhauer's psychology is central to von Hartmann's view, a view that also contains ideas similar to Freud's ideas. Moreover, Brentano discussed von Hartmann at length and on precisely the question of unconscious mental states in his 1874 book.

There is more. First, the period of Freud's secondary schooling and university, roughly from 1865 to 1875, was the period of Schopenhauer's greatest fame. Indeed, he was virtually the official philosopher of the German-speaking world in those years. Most university students in Freud's time who had any interest in philosophy or psychology would have read at least *The World as Will and Representation*. Secondly, as a student, Sulloway (1979) tells us, Freud belonged to *Leseverein der deutschen Studenten Wiens* (Reading Society of the German Students of Vienna). Along with Wagner and Nietzsche, Schopenhauer was a prime topic of conversation. Thirdly, Brentano refers to Schopenhauer several times in his 1874 book.

Schopenhauer and Freud

For all the reasons that we have laid out, we are of the view that the deep similarities in the views of Schopenhauer and Freud were not a coincidence; when Freud said, 'Why should not a bold thinker have guessed something that is afterwards confirmed by sober and painstaking detailed research?' (1933: 107), we suspect that he should have said 'that (p. 78) is afterwards *taken up* and confirmed by sober and painstaking detailed research'.[9] Freud's claim that he made his discoveries without being influenced by Schopenhauer would be like someone today claiming that she or he discovered dynamic psychology without being influenced by Freud!

Freud's ambivalence about Schopenhauer goes with peculiarities in his attitude to philosophers in general. An example. In 1923 and many times in other works, he said that philosophers reject the idea of unconscious psychical states: 'To most people who have been educated in philosophy the idea of anything psychical which is not also conscious is so inconceivable that it seems to them absurd and refutable simply by logic … ' (1923: 13). This is a remarkable statement, considering that Freud was steeped in German Romanticism from his youth. Schopenhauer denied the unconscious? Von Hartmann? Nietzsche? And there were many others. As Ellenberger notes, in the latter part of the nineteenth century, the unconscious was extremely popular (1970: 311).

Not only did this tradition not deny unconscious mental life, it held it to be the most important part of the human person. Freud had to have been aware of all this. So, what is going on? Perhaps something like the following. One philosopher who influenced Freud *did* deny the very possibility of unconscious psychic states: Franz Brentano. (Brentano was not the only nineteenth-century philosopher who equated the psychical with the conscious but the others are not known to have influenced Freud.) Freud attended Brentano's lectures for three years, at just the time Brentano published his famous 1874 book in which the equation of the psychical with the conscious appears most prominently. And Brentano was the only philosopher under whom Freud studied. At the time, as Freud's letter to his friend Silberstein shows, Freud idealized Brentano, describing him at one point as a 'remarkable man and in many respects ideal human being' (1990: 95). Yet Freud mentioned Brentano only once in all his writings. It is possible that when Freud thought of philosophers, he thought of philosophers such as Brentano and when he worked on his own views, he did not think of philosophers such as Schopenhauer and von Hartmann who had held similar views.

We will close by noting that Schopenhauer's contribution likely goes beyond any direct or indirect influence that he had on Freud. Intellectual historians must also consider Schopenhauer's influence on those Freud sought to persuade. In addition to influencing Freud on the many doctrines that we have detailed, it is reasonable to believe that Schopenhauer helped to shape the *reception* of psychoanalysis. Obviously, many of Freud's ideas were shocking to his contemporaries and he struggled against fierce resistance to gain acceptance for them. However, Freud promulgated his theories in an intellectual climate thoroughly shaped by Schopenhauer and other philosophers and psychologists of the unconscious will, Nietzsche for example. Freud himself probably

came to recognize, maybe after 1915, that invoking Schopenhauer helped his cause. If so, Schopenhauer paved the way for Freud in more ways than one.[10]

References

Askey, R. and Farquhar, J (2006). *Apprehending the Inaccessible: Freudian Psychoanalysis and Existential Phenomenology*. Evanston, IL: Northwestern University Press.

Bischler, W. (1939). 'Schopenhauer and Freud: A comparison'. *Psychoanalytic Quarterly, 8*: 88–97.

Brandell, G. (1979). *Freud: A Man of His Century*. New Jersey: Humanities Press.

Brentano, F. (1874/1973). *Psychology from an Empirical Standpoint*, trans. A. C. Rancurello, D. B. Terrell, and L. L. McAlister. London: Routledge and Kegan Paul.

Brook, A. (1992). 'Psychoanalysis and commonsense psychology'. *Annual of Psychoanalysis 20*: 273–303.

Brook, A. (forthcoming). 'Obscure representations'. In J. Wuerth (ed.), *The Cambridge Kant Lexicon*. Cambridge: Cambridge University Press.

Cassirer, E. (1946). *The Myth of the State*. New Haven, CT: Yale University Press.

Ellenberger, H. (1970). *The Discovery of the Unconscious: The History and Evolution of Dynamic Psychiatry*. New York: Basic Books.

Freud, S. (1893–5). *Studies on Hysteria*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 2. London: Hogarth Press.

Freud, S. (1895). 'Project for a scientific psychology'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 1 (pp. 283–410). London: Hogarth Press.

Freud, S. (1900). *The Interpretation of Dreams*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vols 4 and 5. London: Hogarth Press.

Freud, S. (1901). *The Psychopathology of Everyday Life*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 6. London: Hogarth Press.

Freud, S. (1905). 'Three essays on the theory of sexuality'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 7 (pp. 145–245). London: Hogarth Press.

Freud, S. (1909). 'Notes on a case of obsessional neurosis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 10 (pp. 153–326). London: Hogarth Press.

Freud, S. (1910a). 'Five lectures on psycho-analysis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 11 (pp. 3–56). London: Hogarth Press.

Freud, S. (1910b). 'Wild psychoanalysis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 11 (pp. 219–227). London: Hogarth Press.

Freud, S. (1911). 'Formulations on the two principles of mental functioning'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 12 (pp. 215–226). London: Hogarth Press.

Freud, S. (1912). 'A note on the unconscious'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 12 (pp. 257–266). London: Hogarth Press.

Freud, S. (1913). 'The disposition to obsessional neurosis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 12 (pp. 311–326). London: Hogarth Press.

(p. 80) Freud, S. (1914a). 'On the history of the psycho-analytic movement'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 1–71). London: Hogarth Press.

Freud, S. (1914b). 'On narcissism'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 67–102). London: Hogarth Press.

Freud, S. (1915a). 'Instincts and their vicissitudes'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 109–140). London: Hogarth Press.

Freud, S. (1915b). 'Repression'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 141–158). London: Hogarth Press.

Freud, S. (1915c). 'The unconscious'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 159–215). London: Hogarth Press.

Freud, S. (1916–17). *Introductory Lectures on Psycho-Analysis*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vols 15 and 16. London: Hogarth Press.

Freud, S. (1917). 'A difficulty in the path of psycho-analysis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 17 (pp. 135–144). London: Hogarth Press.

Freud, S. (1919). 'Introduction to *Psycho-Analysis and The War Neuroses*'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 17 (pp. 205–215). London: Hogarth Press.

Freud, S. (1920). 'Beyond the pleasure principle'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 18 (pp. 3–64). London: Hogarth Press.

Freud, S. (1923). *The Ego and the Id*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 19 (pp. 3–66). London: Hogarth Press.

Freud, S. (1924a). The Economic Theory of Masochism. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 19. London: Hogarth Press.

Freud, S. (1924b). 'Neurosis and psychosis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 19 (pp. 147–153). London: Hogarth Press.

Freud, S. (1925a). 'An autobiographical study'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 20 (pp. 3–74). London: Hogarth Press.

Freud, S. (1925b). 'Psycho-analysis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 20 (pp. 259–270). London: Hogarth Press.

Freud, S. (1933). 'New introductory lectures on psycho-analysis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 22 (pp. 3–182). London: Hogarth Press.

Freud, S. (1940a [1938]). 'An outline of psycho-analysis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 23 (pp. 139–207). London: Hogarth Press.

(p. 81) Freud, S. (1940b [1938]). 'Some elementary lessons in psycho-analysis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 23 (pp. 279–286). London: Hogarth Press.

Freud, S. (1990). *The Letters of Sigmund Freud to Eduard Silberstein, 1871–1881*, ed. Walter Boehlich, trans. A. Pomerans. Cambridge, MA: Belknap Press of Harvard University Press.

Gardiner, P. (1963). *Schopenhauer*. London: Penguin Books.

Gardner, S. (1999). 'Schopenhauer, Will, and the Unconscious'. In C. Janaway (ed.), *The Cambridge Companion to Schopenhauer* (pp. 375–421). Cambridge: Cambridge University Press.

Gardner, S. (2013). 'The unconscious: Transcendental origins, Idealist metaphysics and psychoanalytic metapsychology'. In N. Boyle and L. Disley, and K. Ameriks (eds), *The Impact of Idealism: The Legacy of Post-Kantian German Thought* (pp. 134–165). Cambridge: Cambridge University Press.

Gödde, G. (2010). 'Freud and 19th-century philosophical sources on the unconscious'. In A. Nicholis and M. Leibscher (eds), *Thinking the Unconscious: 19th-Century German Thought* (pp. 261–286). Cambridge: Cambridge University Press.

Gupta, R. (1980). 'Freud and Schopenhauer'. In M. Fox (ed.), *Schopenhauer: His Philosophical Achievement*. New Jersey: Barnes & Noble Books.

Hartmann, E. von (1869/1931). *Philosophy of the Unconscious: Speculative Results According to the Inductive Method of Physical Science*. New York: Harcourt, Brace & Company.

Herzog, P. (1988). 'The myth of Freud as anti-philosopher'. In *Freud: Appraisals and Reappraisals*, Vol. 2 (pp. 163–190). New Jersey: The Analytic Press.

Jacquette, D. (2015). *Schopenhauer and the Metaphysics of the Unconscious*. Basel: Schwage AG, Verlag.

Janaway, C. (2010). 'The real essence of human beings: Schopenhauer and the unconscious will'. In A. Nicholis and M. Leibscher (eds), *Thinking the Unconscious: 19th-Century German Thought* (pp. 140–155). Cambridge: Cambridge University Press.

Jung, C. (1973). *Memories, Dreams, Reflections*. New York: Vintage Books.

Magee, B. (1989). *The Philosophy of Schopenhauer*. New York: Oxford University Press.

Mann, T. (1968). *Essays of Three Decades*. Toronto: Random House of Canada Ltd.

McGill, V. J. (1971). *Schopenhauer: Pessimist and Pagan*. New York: Haskell.

McGrath, W. (1986). *The Politics of Hysteria: Freud's Discovery of Psychoanalysis*. Ithaca, NY: Cornell University Press.

Proctor-Greg, N. (1956). 'Schopenhauer and Freud'. *Psychoanalytic Quarterly 25*: 197–214.

Schopenhauer, A. (1819/1969). *The World as Will and Representation*, Vol. 1, trans. E. F. J. Payne. New York: Dover Publications.

Schopenhauer and Freud

Schopenhauer, A. (1844/1966). *The World as Will and Representation*, Vol. 2, trans. E. F. J. Payne. New York: Dover Publications.

Schopenhauer, A. (1851/1974). 'Essay on spirit seeing and everything connected therewith'. In *Parerga and Paralipomena*, Vol. 1, trans. E. F. J. Payne. Oxford: Clarendon Press.

Schopenhauer, A. (1970). *Essays and Aphorisms* (selection and translation of *Parerga and Paralipomena*, 1851). London: Penguin Books.

Sulloway, F. (1979). *Freud, Biologist of the Mind: Beyond the Psychoanalytic Legend*. New York: Basic Books.

Young, C. and Brook, A. (1994). 'Schopenhauer and Freud'. The *International Journal of Psychoanalysis 75*: 101–118. (p. 82)

Notes:

([1]) In Young and Brook (1994), we explore earlier studies: Cassirer (1946); Ellenberger (1970); Gupta (1980); Gardiner (1963); Magee (1989); and Mann (1968). More recent studies include: Gardner (1999, 2013); Askey and Farquhar (2006); Janaway (2010); and Jacquette (2015). Gödde (2010) provides an excellent synopsis of the whole period. Sandford (this volume) paints a picture of Schopenhauer and Freud on reproduction and sexuality that is different from ours. None of these authors goes into anything like the detail that we do.

([2]) We will limit ourselves to correspondences in psychological doctrine between the two thinkers, though there are other similarities in their views, too, for example in their ethics and aesthetics.

([3]) Though *The World as Will and Representation* was published in 1819, we date Vol. 2 to 1844 because it did not appear until the second edition, which was published in 1844. A third edition was also published in Schopenhauer's lifetime and he expanded the work still further. As he tells us in the Preface, 'though in the same type', it has '136 pages more than its predecessor'.

([4]) Sandford (this volume) argues that Schopenhauer was concerned with a desire to reproduce, whereas Freud was concerned with sexual desire. Statements such as the ones just quoted suggest that she pushes her argument too far, and that while Schopenhauer certainly linked sexual desire to the will to live and reproduce (sexual desire 'is the most complete manifestation of the will to live' (1844 (2): 514)), by sexual desire he often had plain old lust in mind.

([5]) Freud's expansion of the concept of sexuality is altogether more complicated than Schopenhauer's. We explore some of the complications in Young and Brook (1994).

(⁶) In response, we might try to say that the sensual zones and pleasures of Freud's theory of infantile sexuality are sexual only in the extended sense of the term we delineated earlier. Freud no more believed that infantile sexuality is aimed at orgasmic release or that it is genital than anyone else—indeed, his stages of psychosexual development were specifically built on a denial of the latter point. If so, it may sometimes appear as if his infantile 'sexuality' comes down to little more than 'organ-pleasure', bodily sensuality in general (see in this regard the argumentative discussion at 1916–17: 323–5). Such a minimalist reading would seriously understate the originality of his theory, however; for Freud, the infantile part-instincts and erogenous zones are sexual in a much stronger sense than that—they are the *origins* of genital sexuality in the human organism.

(⁷) Kant held much the same view too (Gardner 2013: 136).

(⁸) Freud uses the word *Widerstand* for resistance, Schopenhauer here uses *Widerstreben*.

(⁹) Moreover, Schopenhauer may have been less a mere speculator than Freud alleges. As Magee notes, he 'was a frequent visitor to insane asylums, where he would hold long conversations with the inmates, and go back again and again to talk to those who particularly interested him' (1988: 266).

(¹⁰) An earlier version of this paper appeared as Brook & Young (1994).

Andrew Brook

Department of Philosophy, Carleton University, Ottawa, Canada

Chris Young

Independent scholar

Oxford Handbooks Online

Abstract and Keywords

The notion of sublimation is essential to Nietzsche and Freud. However, Freud's writings fail to provide a persuasive notion of sublimation. In particular, Freud's writings are confused on the distinction between pathological symptoms and sublimations, and on the relation between sublimation and repression. After rehearsing these problems in some detail it is proposed that a return to Nietzsche allows for a more coherent account of sublimation, its difference from pathological symptoms and its relation to repression. In summary, on Nietzsche's account, while repression and pathological symptoms involve a disintegration (of the self), sublimation involves integration. The chapter provides a brief consideration of some post-Freudian accounts of sublimation that, arguably, represent a return to a more Nietzschean approach. In closing, the chapter contrasts the different bases of Nietzsche's and Freud's valorization of sublimation.

Keywords: sublimation, repression, pathological symptoms, integration, disintegration, master drive, Shreber, Leonardo da Vinci

Introduction

Jacob Golomb writes that '(f)or Freud, sublimation is the only alternative to excessive repression which may deteriorate into an acute neurotic conflict'.[1] He quotes Freud:

> It is precisely in neurotics that … the formation of an ideal heightens the demands of the ego and is met with the most powerful factor favouring repression; sublimation is a way out, by which those demands can be met without involving repression
>
> (Freud 1914: 95).

He then adds: 'This comment accurately reflects the ideas that Nietzsche had formulated on sublimation years before Freud' (Golomb 1989: 69–70). Golomb argues correctly that sublimation is indeed a central notion for both Nietzsche and Freud, but as we shall see, the question of what exactly is involved in sublimation raises profound difficulties. The penultimate sentence in the entry on sublimation in Laplanche's and Pontalis's canonical *The Language of Psychoanalysis* explicitly acknowledges the importance of giving a coherent theory of sublimation:

> In the psychoanalytic literature the concept of sublimation is frequently called upon; the idea answers to a basic need of the Freudian doctrine and it is hard to see how it could be dispensed with
>
> (Laplanche and Pontalis 1973: 433).

One reason that sublimation is a key notion in psychoanalysis is that from Freud's therapeutic point of view successful psychoanalytic treatment aims at sublimation, in as much as sublimation is seen as a necessary condition for psychic health. By bringing to conscious light hitherto repressed drives, desires, and wishes, energy which has previously displayed itself in unpleasurable symptoms may be harnessed and directed to more productive and felicitous ends. And indeed, at first glance, sublimation might seem a clear enough concept. It involves the redirecting of a repressed sexual drive towards a non-sexual aim.[2] As Freud puts it in his paper *On Narcissism*:

> Sublimation is a process that concerns object-libido and consists in the instinct directing itself towards an aim other than, and remote from, that of sexual satisfaction; in this process the accent falls upon deflection from sexuality[3]
>
> (Freud 1914: 94).

Yet Laplanche's and Pontalis's entry on sublimation ends with a starkly negative assessment of attempts to clarify the notion of sublimation: 'The lack of a coherent theory of sublimation remains one of the lacunae in psychoanalytic thought' (Laplanche and Pontalis 1973: 433).

The claim that sublimation is a vexed concept is repeatedly echoed in the 'secondary' literature. Indeed, this is true from the early days of that literature as evidenced in Edward Glover's (1931) assertion in his important paper 'Sublimation, Substitution and Social Anxiety':

> It is generally agreed that prior to 1923 a good deal of confusion existed regarding the exact nature of sublimation. Since then it has increased rather than diminished
>
> (Glover 1931: 263),

More recently Jacques Lacan notes 'the virtually absurd difficulties that authors have encountered every time they have tried to give a meaning to the term "sublimation"' (Lacan 1992: 142–3).

To get an entry to these difficulties let us begin by considering the simple definition given here which emphasizes the redirecting of a drive from a sexual to a non-sexual aim. Before demonstrating the major problem with this definition we shall briefly consider two minor but important problems with this definition. First, there is the problem of distinguishing aims that involve sexual satisfaction and those that do not. After all, it was Freud who argued that many prima facie non-sexual activities (e.g. thumb-sucking) have a sexual component. Perhaps this can be finessed by distinguishing overtly sexual aims from aims that are not overtly sexual. Second, there is the problem that much sublimation seems to involve a diversion of arguably non-sexual instincts. Thus aggressive drives, drives associated with what Freud would later call the death drive, may find sublimated discharge in non-aggressive behaviour.[4]

Substitute Formations, Symptoms, and Sublimations

Now for the major problem: the definition in terms of the substitution of non-sexual for sexual aims fails to distinguish sublimations from symptom formation. An obsessive who compulsively aims to avoid treading on cracks in the pavement may have substituted a non-sexual aim for a sexual aim but that would not normally count as a case of sublimation (though see the following section). On Freud's basic account various drives are repressed and later these may manifest themselves in various behaviours. These behaviours, often labelled by Freud as 'substitute formations', are related in complex ways to the original drive. They are what Freud calls the 'return of the repressed' (Freud 1915: 154). Our key question then is how to separate those substitute formations, those instances of the return of the repressed, which are symptoms, from those that are sublimations?

Of all Freud's writings perhaps the paper that most acutely demonstrates the need for a clear means of separating symptoms from sublimations is his 1910 essay 'Leonardo Da Vinci and a Memory of his Childhood' (Freud 1910a). There Freud argues that Leonardo is a model of that type of individual whose repressed sexual desires find sublimated expression by becoming attached to a powerful drive for scientific and artistic research. In his summarizing final section Freud shows implicit recognition of the difficulties of separating symptoms from sublimations, and, relatedly, illness from health, when he writes:

> We must expressly insist that we have never reckoned Leonardo as a neurotic or a 'nerve case', as the awkward phrase goes. Anyone who protests at our so much as daring to examine him in the light of discoveries gained in the field of pathology is still clinging to prejudices which we have to-day rightly abandoned. We no longer

think that health and illness, normal and neurotic people, are to be sharply distinguished from each other, and that neurotic traits must necessarily be taken as proofs of a general inferiority. To-day we know that neurotic symptoms are structures which are substitutes for certain achievements of repression that we have to carry out in the course of our development from a child to a civilized human being. We know too that we all produce such substitutive structures, and that it is only their number, intensity and distribution which justify us in using the practical concept of illness and in inferring the presence of constitutional inferiority. From the slight indications we have about Leonardo's personality we should be inclined to place him close to the type of neurotic that we describe as 'obsessional'; and we may compare his researches to the 'obsessive brooding' of neurotics, and his inhibitions to what are known as their 'abulias'

(Freud 1910a: 131).

After holding up Leonardo as a model of sublimation throughout the essay, and then in the first sentence of this quotation protesting that he has not reckoned Leonardo to be a neurotic, Freud in the final sentence concludes that we must place Leonardo close to the obsessional neurotic. The distance between sublimation and neurotic symptoms seems vanishingly small. Here we have in microcosm our problem; how to separate sublimations from the neurotic symptoms both of which count as substitutive formations?

The Social Factor in Sublimations

In many of Freud's explications of sublimation there is an emphasis on the social value of the activity resulting from sublimation. Thus in his *New Introductory Lectures* he says:

A certain kind of modification of the aim and a change of object, in which our social valuation is taken into account, is described by us as 'sublimation'

(Freud 1933: 97).

In one of his earliest published references to sublimation in the Dora case of 1905 he talks of the 'the undifferentiated sexual disposition of the child ... being diverted to a higher sexual aim' and thereby providing 'the energy for a greater number of our cultural achievements' (Freud 1905a: 50). In the *Three Essays* dating from the same period he writes that sexual curiosity can be 'diverted (sublimated) in the direction of art' (Freud 1905a:156).

This emphasis on socially valued achievements would provide some means of separating neurotic symptoms from sublimations but at the high cost of introducing a totally non-psychoanalytic, indeed a non-psychological element to the definition; namely that of social valuation. One, presumably uncomfortable, consequence of such an account would be that as social values change, behaviour that was at one time pathological would become a sublimation. For instance, suppose our compulsive crack avoider mentioned earlier successfully presents his activities as a kind of performance art and receives much

social acclaim for his continual performances. Does that move his activity from neurotic symptom to sublimation?

The reference to social acclaim has led many authors to be suspicious of such definitions.[5] In his article 'What is Sublimated in Sublimation' Donald Kaplan begins:

> The whole idea of sublimation has been a vagrant problem for psychoanalysis from the very beginning. This is because sublimation, in one of its meanings, refers to felicitous exercises of the mind, while psychoanalysis is suspicious of any such simple creed

> (Kaplan 1993: 549).

One gets a clear picture of the problem posed by the introduction of social valuation as a defining characteristic of sublimation, and the related problem of distinguishing neurotic symptoms from sublimations, by contrasting Freud's essay on Schreber with his essay on Leonardo. The latter is perhaps his most sustained surviving piece covering the topic of sublimation.[6] According to Freud, Schreber's illness stems from his repression of homosexual desires first overtly expressed in his quickly suppressed thought that '(a)fter all it would be nice to be a woman submitting to the act of copulation'. This repression later gave rise to the obsessive thought that God is trying to unman him and turn him into a woman and led to various other obsessive thoughts and behaviour. This would not normally count as a case of sublimation. Contrast this with Freud's Leonardo case. According to Freud, Leonardo, like Schreber, represses his homosexual desires and this repression later leads to his scientific inquisitiveness and inventiveness and his artistic activities including his obsession with capturing in drawings perfect representations of the male body. This for Freud is a classical case of sublimation. Is the only pertinent difference here that in the Schreber case the end activity that results from a repression of erotic desires, namely, obsessive behaviour and thought centring on his relationship to God, is not considered socially valuable, whereas in the Leonardo case the end activity that results from a repression of erotic desires is considered socially valuable? Or, as Kaplan might put it, is it simply that we regard Schreber's exercises of the mind as non-felicitous, but regard Leonardo's as felicitous? Such distinctions as socially valuable/not socially valuable, felicitous/non-felicitous, clearly involve value judgements. Moreover, such distinctions do not seem to be analytical distinctions. Thus it seems a fundamental notion in psychoanalysis is not to be explained in strictly psychoanalytical terms. Furthermore, those terms seem irredeemably normative and this runs against Freud's general claim to be providing a merely scientific/descriptive account of psychological phenomena.[7] Also bringing in the normative results in a questionable form of relativism as norms of what is socially acceptable vary from culture to culture and even between subgroups in a given culture. It is presumably this kind of concern that lead Laplanche and Pontalis to raise the question: 'Should the fact that activities described as sublimated in a given culture are accorded particularly high social esteem be taken as a defining characteristic of sublimation?' (Laplanche and Pontalis 1973: 33).

If we do choose to abandon the social valuation factor in separating sublimation from symptom formation any new account should, I suggest, be tested by how it handles the Leonardo and Schreber cases. The respective behaviours of Leonardo and Schreber may be taken as paradigms of sublimation and pathogenic symptoms. What we are looking for then is a strictly psychoanalytic account which would class Leonardo's behaviour as sublimation and Schreber's as pathogenic symptoms.

The Relation between Sublimation and Repression

Another serious problem raised by Freud's own writings on sublimation concerns the relation between repression and sublimation. In one understanding of Freud's telling of the Leonardo case, Leonardo's homosexual urges are continuously repressed, hence he lives an asexual adult life, and it is this repression that gives rise to his adult scientific and artistic drives:

> the sexual repression which set in after this phase of his childhood (where he experienced the 'excessive tenderness of his mother') caused him to sublimate his libido into the urge to know and established his sexual inactivity for the whole of his later life

> (Freud 1910a: 135).

Here it seems out of place to talk of the repression of the original drives being lifted; for if the repression were lifted there would be no account of the continuation of the scientific and artistic drives and the continual asexuality.[8] On this model, at its simplest, Leonardo's scientific and artistic drives, like his asexuality, are really kinds of symptoms and, as per the usual Freudian model, working through the repressions behind the symptoms should lead to elimination, or possibly the transformation, of the symptoms. The notion that sublimation involves repression is re-enforced by one of Freud's earliest mentions of sublimation in his *Three Essays on Sexuality*.[9] There he refers to 'the diversion of sexual forces from sexual aims and their direction to new ones—a process that deserves the name of "sublimation"' (Freud 1905a: 178). In the same place he goes on to say 'we would place its (sublimation's) beginning in the period of sexual latency of childhood'. Clearly such redirection on the part of the child involves non-conscious repression of those sexual forces qua sexual forces, and such repression is exactly what Freud describes as the progenitor of the period of sexual latency. It is not that the child consciously faces his erotic desires and chooses not to act on them but pursue some other activity as a substitute satisfaction. The sublimation of those desires first requires their repression. Indeed, there is even some tendency in the psychoanalytic literature to at least implicitly identify sublimation as a species of repression. For instance in *Sublimation: Inquiries into Theoretical Psychoanalysis*, Hans Loewald writes 'Traditionally sublimation is classified in psychoanalysis as a defence' (Loewald 1988: 3), and in fact Freud's earliest reference to sublimation in his 1892 letter to Fliess refers to

sublimations as 'protective structures' (Freud 1897: 247). A page later in a footnote Loewald notes that:

> It must be noted, however, that in the psychoanalytic literature, including works by Freud himself, the very distinction between defence and repression has been blurred, in as much as repression is used interchangeably with defense
>
> (Loewald 1988: 4–5).

So if sublimations are defences and the notions of defences and repressions are interchangeable it follows that sublimations are repressions. Fenichel indeed comments that sometimes sublimation is called a 'successful repression' (Fenichel 1945: 148).

In fact while the Leonardo text and the extract from the *Three Essays on Sexuality* suggest that repression is part of sublimation, often in other texts when Freud explicitly considers repression and sublimation together he claims that sublimation is an alternative to repression. Thus in a passage from *Five Lectures on Psychoanalysis* he says: 'Premature repression makes the sublimation of the repressed instinct impossible; when the repression is lifted, the path to sublimation becomes free once more' (Freud 1910b: 54). In a letter to Putnam, Freud even goes so far as to claim that the two are mutually exclusive: 'It (psychoanalytic theory) teaches that a drive cannot be sublimated as long as it is repressed and that is true for every component of a drive' (Hale 1971: 121).

To clarify the relation between sublimation and repression it helps to consider exactly what gets repressed in repression. Here the best place to turn is Freud's (1915) essay 'Repression' (Freud 1915: 141–58). A crucial point of that essay is Freud's claim that a drive has both an 'ideational' component and an energetic component, what Freud calls 'a quota of affect' (Freud 1915: 152). The ideational component is the intensional content of the drive, including its aim. The energetic component is the force associated with that aim, including the strength of the drive. In repression both the ideational component and the force component are repressed. The repression of the ideational component involves not letting the aim be consciously apprehended. The repression of the force involves not letting the force be expressed in behaviour. While Freud refers to the repression of the ideational component as 'primal repression' he emphasizes that 'the vicissitude of the quota of affect belonging to the representative is far more important than the vicissitude of the idea' (Freud 1915: 153). With this distinction in hand we can now see how sublimation both does and does not involve repression. In sublimation the ideational component may or may not be repressed. In the case of childhood sublimations, the aim of the original drive, the drive's ideational component, is typically held back from conscious apprehension. In the case of sublimation reached through psychoanalytic treatments typically the ideational component becomes available for conscious apprehension. However, in all sublimations, therapeutically achieved or otherwise, the force component is expressed in behaviour. All sublimations involve an expression of a pent-up quota of affect. The picture suggested here is that all sublimations typically take repressions as causal antecedents. In this sense sublimations are another manifestation

of the phenomena that Freud calls 'the return of the repressed'. What sublimations undo is the repressing of the energetic component, they steer it to an outlet, an aim that deviates from its original aim. Indeed in the passage quoted earlier from 'Five Lectures on Psychoanalysis' Freud contrasts neurotics who owing to their repressions 'have sacrificed many sources of energy' with cases of sublimation:

> in which the energy of the infantile wishful impulses is not cut off but remains ready for use—the unserviceable aim of the various impulses being replaced by one that is higher, and perhaps no longer sexual
>
> (Freud 1910b: 53–4).

Here Freud is explicitly claiming that in sublimation it is the repression of the energy that is lifted.

This reading does not seem to square with the letter to Putnam where Freud says that sublimation precludes the repression of *any* component of a drive. However, it is worth noting that in that letter he is focused on the clinical practice of psychotherapy. Presumably his idea is that the analyst is faced with a patient presenting various symptoms which the patient experiences as distressing. These pathogenic symptoms are the manifestation of various unconscious repressions. By bringing the repressed drives into full consciousness—this of course involves more than just the patient's intellectual recognition of the drives—this allows for their expression in more acceptable forms. Note, on this model it is not that the ideational component of the drive ceases to be *suppressed* but that the suppression takes on a conscious form that somehow allows the energetic component formerly associated with it to be redirected to new, more acceptable ends. This model does not fit the non-therapeutic cases such as that of Leonardo since there is no conscious recognition on Leonardo's part of his underlying homoerotic drives. Nor does it fit the text of Freud's *Three Essays on Sexuality* where the child in his latency period has no conscious awareness of the desires being sublimated. Clearly the child (unconsciously) represses rather than (consciously) suppresses his erotic desires. So the conclusion to be drawn here is that the sublimation that is the result of successful psychoanalytic intervention involves the lifting of unconscious repression of both the ideational and energetic component; but the sublimation involved in typical cases of scientific and artistic expression or that occurs in the latency period in childhood typically only involve lifting of the repression of the energetic component and continued repression of the ideational component. I suspect it is this model that leads to Fenichel's idea of sublimations as successful repressions; they are successful in that while the ideational component is kept from consciousness the energy associated with it is allowed an outlet. Freud's claim in the Putnam letter that a drive cannot be sublimated as long as any component is repressed would then have to be read as asserting that 'in the practice of psychotherapy a drive cannot be sublimated as long as it is unconsciously repressed'.

Nevertheless, this understanding of the relation of repression and sublimation still leaves us hard-pressed to distinguish sublimation from other instances of the return of the repressed, such as pathological symptoms.

Nietzsche's Solution

Nietzsche, like Freud[10] and like their common predecessor Schopenhauer, takes individual humans to be, at some fundamental level, collections of drives.[11] However, most modern humans, as members of what he denigratingly calls the herd, are simply disorganized collections of competing drives, with different drives having relative ascendancy at different times:

> In the present age human beings have in their bodies the heritage of multiple origins, that is opposite and not merely opposite drives and value standards that fight each other and rarely permit each other any rest. Such human beings of late cultures and refracted lights will on the average be weaker human beings (BGE 200).[12]

In many cases drives, particularly aggressive drives, are treated as per the Freudian model by repression:

> All instincts which are not discharged to the outside are turned back inside. This is what I call the internalization of man. From this first grows in man what people later call his 'soul'. The entire inner world, originally as thin as if stretched between two layers of skin, expanded and extended itself, acquired depth, width, and height to the extent that the discharge of human instinct out into the world was obstructed (GM II 16).

Central to Nietzsche's account is the notion of splitting off. Aggressive drives, which are not viewed as acceptable, typically because acting on them would exact a painful retribution, are suppressed to the point that one does not even acknowledge that one has them; that is to the point where they are not simply suppressed but repressed.[13] They may nevertheless find their outlet often in disguised form. Indeed, often in a form that contradicts their very nature. Thus Nietzsche claims that the Christian value of brotherly love was originally, in fact, a transformed expression of hostile drives to dominate one's fellow men. By successfully preaching brotherly love the weak get their oppressors to voluntarily disarm themselves and become subservient to the values of the weak. In doing so they, both the weak and the strong who have been converted to the values of the weak, split off their contrary aggressive drives from conscious apprehension, so that at the same time they harbour both unacknowledged aggressive drives and acknowledged beneficent drives. This repression and splitting-off of drives makes us sick creatures of what Nietzsche calls *ressentiment*.[14]

For Nietzsche in certain rare cases drives, rather than being split off, are harnessed into a centred, unified, whole. At one point Nietzsche took Wagner to be such a case:

> The dramatic element in Wagner's development is quite unmistakable from the moment when his ruling passion became aware of itself and took his nature in its charge: from that time on there was an end to fumbling, straying, to the proliferation of secondary shoots, and within the most convoluted courses and

often daring trajectories assumed by his artistic plans there rules a single inner law, a will by which they can be explained (UM III 2).

The story of Wagner's achievement of a higher unity borne from some master drive is of course the story Nietzsche would repeat about himself in the dramatic section of *Ecce Homo* where Nietzsche elaborates the subtitle of that work 'How One Becomes What One Is':

> To become what one is, one must not have the slightest notion of what one is … The whole surface of consciousness—consciousness is a surface—must be kept clear of all great imperatives … Meanwhile the organizing 'idea' that is destined to rule keeps growing deep down—it begins to command; slowly it leads us back from side roads and wrong roads; it prepares single qualities and fitnesses that will one day prove to be indispensable as a means towards the whole—one by one, it trains all subservient capacities before giving any hint of the dominant task, 'goal', 'aim', or 'meaning' (EH, Why I am So Clever, 9).

This notion of the training of subservient drives is to be explicated in terms of the redirection of those drives away from their initial, primary, goal towards a secondary goal that is more in line with the master drive. This idea is partially expressed in the following passage from *Human, All too Human*:

> *Microcosm and macrocosm of culture*. Man makes the best discoveries about culture within himself when he finds two heterogeneous powers governing there. Given that a man loved the plastic arts or music as much as he was moved by the spirit of science, and that he deemed it impossible to end this contradiction by destroying the one and completely unleashing the other power; then, the only thing remaining to him is to make such a large edifice of culture out of himself that both powers can live there, even if at different ends of it; between them are sheltered conciliatory central powers, with the dominating strength to settle, if need be, any quarrels that break out. Such a cultural edifice in the single individual will have the greatest similarity to the cultural architecture of whole eras and, by analogy, provide continuous instruction about them. For wherever the great architecture of culture developed, it was its task to force opposing forces into harmony through an overwhelming aggregation of the remaining, less incompatible powers, yet without suppressing or shackling them (HAH I, 276).

A point to be emphasized here is that on Nietzsche's ideal weaker drives are not suppressed or shackled. Rather they are to be harnessed to allow their expression in service to a higher aim. Thus in his note books Nietzsche writes:

> Overcoming of the affects? No, if that means their weakening and annihilation. But instead employing them; which may mean a long tyrannizing of them … At last they are confidently given freedom again: they love us as good servants and

happily go wherever our best interests lie (KSA 12:39, see also HAH I, 224 and 276).

This gives us the material we need to affect a Nietzschean account of the distinction between repression and sublimation. Sublimation is what happens when a drive's primary aim is substituted by a secondary aim that allows for expression of the drive in a manner consonant with the master drive. As John Richardson succinctly puts it: '(D)rive A rules B insofar as it has turned B towards A's own end, so that B now participates in A's distinctive activity' (1996: 33).[15]

Repression is what happens when a drive is denied its immediate aim and is then split off from other drives in the sense that its aims are not integrated with the aims of other drives and it must battle, often unsuccessfully, for any opportunity to achieve expression.

Consider again our cases of Leonardo and Schreber. Leonardo's homosexual sexual drive is redirected towards the secondary aims of scientific, as opposed to sexual, researches and artistic creation, including idealized representations of the male body, as opposed to actual sexual possession of male bodies. These ends fit in with his master drive of scientific and artistic creativity. Schreber, on the other hand, by his own admission, splits off his religious activity completely from his life as senate president. He claims that neither activity in a way informs the other. Schreber does not integrate his sexual drive with his wider life but strives to isolate it and separate its expressions from his other activities.[16]

The Nietzschean solution to the problem of differentiating sublimation from pathological symptoms may be summed up in the slogan *sublimations involve integration or unification, while pathological symptoms involve splittingoff or disintegration*, as we might call it. What is disintegrated is of course the (unified) self. For Nietzsche the difference between repression and sublimation is that in sublimation the stronger drive co-opts a weaker drive as an ally and this allows the weaker drive expression albeit to an end that contains some degree of deflection from its original aim (for example Leonardo's homosexual drive is expressed in possession of idealized representations of male bodies rather than literal possession of male bodies); whereas in repression the stronger drive attempts to stifle any expression of the weaker drive so that its expression is either fully stifled or can only be achieved in a heavily disguised form which often represents the inverse of the original aim (the Christian's hate and envy of, and desire to have power over his fellow man, being expressed as professions of brotherly love and disinterest in power). The relation between sublimation and repression is that often but not always sublimations have repressions as antecedents. If a stronger drive does not at first have sufficient strength to co-opt a weaker drive to it own ends it may simply act to stifle that weaker drive.[17] As the stronger drive gains in strength opportunities for co-option rather than stifling may arise.[18]

Sublimation and *Ressentiment*

For Nietzsche unified selves, what he takes to be genuine persons, are rare achievements; hence he cautions 'one should not at all assume that many humans are "people"'(KSA 12: 491, translation mine), and '(m)ost men present pieces and fragments of man: one has to add them up for a complete man to appear' (KSA 12: 519).[19] It is Nietzsche's aim to foster the development of such genuine persons. In the same vein his Zarathustra says 'it is my art and aim, to compose into one and bring together what is fragment and riddle and dreadful chance' (TSZ II, 21). Sublimation, for Nietzsche, is the primary means to a unified self. Nietzsche, like Freud, takes sublimation as a mark of health.

Nietzsche's important and difficult normative ideas of *amor fati*, eternal recurrence, and affirmation of life are all strongly related to his aim of overcoming *ressentiment*.[20] *Ressentiment* is directly connected to repression in that where there *is ressentiment* there is some drive that we have been forced to stifle. Nietzsche claims that in order to fully love fate or to fully wish back everything eternally, both of which are exemplary ways of affirming life, we would have to overcome all such *ressentiment*. To affirm all of one's life, to overcome *ressentiment*, would be to affirm all of one's drives; life, for Nietzsche, being nothing but a collection of drives.[21] This does not mean to simply let all of one's drives have free expression. That would involve conflict, chaos and, inevitably, disintegration.[22] It means harnessing one's drives to allow them a form of concerted expression, as Nietzsche himself nicely puts the point:

> The multitude and disgregation of impulses and the lack of any systematic order among them results in a 'weak will'; their coordination under a single predominant impulse results in a 'strong' will: in the first case it is the oscillation and lack of gravity; in the later, the precision and clarity of direction (KSA 13: 394).

Sublimation is for Nietzsche the key means to such concerted expression and, hence, to overcoming *ressentiment*.[23]

On Nietzsche's picture of sublimation the questionable element of social valuation drops out of the picture.[24] Another important difference from the Freudian account is that the notion of desexualization also drops out. Where in his account of sublimation Freud deals primarily with a libidinal drive (Eros), Nietzsche allows for a multiplicity of drives. Relatedly, for Nietzsche aggressive drives as well as sexual drives are prone to sublimation. Given the problems of Freud's notion of desexualization, and his concomitant failure to revise his account of sublimation after his positing of a death drive separate from the libidinal drive, these differences in Nietzsche's account are a gain rather than a loss. The price paid for this gain is that Nietzsche, much more than Freud, remains unclear about the range of drives available for sublimation and repression.[25]

On Nietzsche's account the role of the ego in repressions and sublimations is dramatically reduced. Rather than having a higher level agency, such as Freud's ego or superego censor lower level drives, Nietzsche simply posits an agonal struggle between the drives; thus he says in *Beyond Good and Evil*: 'The will to overcome an affect is, in the end, itself only the will of another, or several other, affects' (BGE 117). And even in those cases where there is consciousness of the drive that is causing a disturbance, that consciousness is not the engine of repression:

> (*T*)hat one *wants* to combat the vehemence of a drive at all, however, does not stand within our own power; nor does the choice of any particular method; nor does the success or failure of this method. What is clearly the case is that in this entire procedure our intellect is only the blind instrument of *another drive*, which is a *rival* of the drive whose vehemence is tormenting us … While 'we' believe we are complaining about the vehemence of a drive, at bottom it is one drive *which is complaining about the other*; that is to say: for us to become aware that we are suffering from the *vehemence* of a drive presupposes the existence of another equally vehement or even more vehement drive, and that a *struggle* is in prospect in which our intellect is going to have to take sides (D: 109).

While this passage allows that our conscious selves may be aware of the drive that is causing us distress, it emphasizes another drive, not the conscious I, as the repressing force. Other texts from Nietzsche make clear that he envisages such struggles between drives often occurring with no conscious awareness, that is with no participation of the conscious I, or ego, as Freud calls it (cf. for instance, the passages quoted earlier describing Wagner and Nietzsche himself).

Klein, Segal, and Loewald: Towards a Reconciliation of Freud and Nietzsche

While most of Freud's texts on sublimation emphasize notions such as desexualization and social valuations, there are occasional texts that emphasize the unifying element in sublimations. In *The Ego and the Id* Freud tells us that sublimated energy may 'retain the main purpose of Eros—that of uniting and binding—as it helps towards establishing unity, or a tendency to unity' (Freud 1923: 45).[26] While Freud fails to thematize this notion of sublimation as unification, some post-Freudians, in particular those working in the tradition of Melanie Klein, do attempt to thematize the important notions of disintegration and integration in their discussions of sublimation.

In her essay 'Infant Analysis' Klein explicitly turns to consideration of Freud's analysis of Leonardo in order to 'set forth more exactly the analogies and differences between symptoms and sublimations' (Klein 1926: 41). The key component she attributes to Leonardo which allows him to achieve sublimations rather than hysterical conversions is his capacity for 'identification with the objects of the world around him'. It is this which allows his fixations to 'become consonant with the ego' (Klein 1926: 43). According to

Klein, having introjected various objects which are assigned both good and bad qualities —this for Klein is a feature of what she calls the depressive position—sublimation is achieved when hostility to introjected objects is overcome through phantasies that allow a symbolic representation of the introjected object in a non-hostile form. It is through such phantasies and symbolic representations that these introjected objects are integrated with the ego. While we cannot hope to fully rehearse here Klein's notion of depressive position etc. what is clear is that for Klein sublimation involves integration and overcoming of hostility to inner objects. Since these inner objects are for Klein manifestations of drives (for instance bad objects are manifestations of aggressive drives), Klein's account of sublimation bears a striking resemblance to the Nietzschean account offered earlier.[27]

The Kleinian account of sublimation as integration is a theme taken up by Hanna Segal in her 'A Psycho-Analytical Approach to Aesthetics' which concludes that 'a satisfactory work of art is achieved by a realization and sublimation of the depressive position' (Segal 1952: 206). For Segal this sublimation involves 'the desire to unite' (Segal 1952: 207). Art, for Segal is a form of sublimatory activity, a working through of the depressive position, that allows us through symbolic representations to take objects towards which we have sadistic hostile impulses and reintegrate them into a world which is 'whole, complete and unified' (Segal 1952: 204).

Another psychoanalyst, influenced by Klein, who identifies sublimation with a type of integration, albeit under the name of 'internalization', is Hans Loewald. Internalization according to Loewald involves 'unconscious ego processes that undo the splitting off' that is characteristic of repression (Loewald 1973: 12). The splitting off that characterizes repression, according to Loewald, 'maintains psychic processes and structures on lower organizational levels' while sublimation, or internalization as he calls it, 'leads to higher organization and an enriched psychic life' (1973: 14). In his later book *Sublimation*, in reference to Freud's Leonardo case, Loewald notes that 'there are unitary experiences which give way to experiences of differentiation, but that in sublimation the experience of unity is restored' (Loewald 1988: 45). Interestingly, Loewald (1973) cites Nietzsche, and in particular Nietzsche's *Genealogy of Morals*, as a precursor for his own notion of internalization. In fact Nietzsche's term 'internalization' is closer to the notion of repression than sublimation. For Nietzsche, as expressed in the passage from GM II 16 quoted earlier, internalization is the turning back of instincts that cannot be discharged externally.

What is missing from these post-Freudian accounts of sublimation as integration or unification, or a higher level of organization, is an account of what exactly is meant by integration, unification, or higher organization. Nietzsche's account of a master drive with a determinate aim realigning the aims of weaker drives, to aims which augment rather than conflict with the aim of the master drive, at least provides a start to such an account. Whether Kleinians or other post-Freudians can utilize such a Nietzschean conception of sublimation as unification without joining Nietzsche in putting such strong conditions on what is necessary for unification, or without joining Nietzsche in his

rejection of the notion of the role of the ego, is something that awaits further investigation.

The Value of Sublimation

Above it was mentioned that both Nietzsche and Freud value health and took sublimation to be a means toward health. Yet where Freud's notion of health involves a fairly moderate normative agenda, Nietzsche's notion of health involves a much more radical normative agenda. Freud, working more within a traditional medical model, took his therapy to be primarily aimed at relieving psychic distress. In this vein he spoke of 'transforming … hysterical misery into common unhappiness' (Freud 1893–5: 305). For Freud, repression leads to anxiety—in his quasi-energetic model this is expressed in terms of the affective state caused by inhibition of discharge; pent up energy denied realization causes a kind of system instability. Here it is clear why, given the aim of a reduction of misery, sublimation, which allows for discharge of the drives, is to be preferred to repression. Continuing in terms of the energetic model, repression leads to a building up of pressure, sublimation allows for a release of pressure.

However on Nietzsche's evaluative scheme the preference for sublimation over repression is less clearly grounded. Nietzsche certainly did not place a high value on contentment and often inveighed against it as a form of complacency; what he disparagingly refers to as 'green pasture happiness of the herd' (BGE 44)—implying that such a goal is worthy only of bovine animals. Where Nietzsche places the pursuit of happiness as the aim of the last men and other lower types (Z Prologue 5), Freud, to some extent echoing Schopenhauer, postulates it as the aim of humans in general:

> They strive for happiness; they want to become happy and to remain so. This endeavour has two sides, a positive and a negative aim. It aims, on the one hand, at an absence of pain and unpleasure, and, on the other, at the experiencing of strong feelings of pleasure
>
> (*Civilization and Its Discontent*, Freud 1930: 76).

Against this type of characterization Nietzsche, presumably having in mind various utilitarian philosophers, quipped: 'Man does not strive after happiness; only the Englishman does that' (TI, Maxims and Arrows, 12).

In *Zarathustra* Nietzsche refers contemptuously to 'the last man' as the 'most contemptible man' who has 'discovered happiness' and who 'makes everything small'; such a man 'will no more shoot the arrow of his longing out over mankind and the string of his bow will have forgotten how to twang' (Z, Prologue, 5). This metaphor of the dissipation of tension contrasts vividly with his characterization of highly repressed individuals who seethe with ressentiment. Indeed one of the most startling of his characterizations of the generally negatively viewed psychic phenomena of ressentiment is that it helped make man an 'interesting animal', that it gave him depth and that

without it '(h)uman history would be much too stupid an affair' (GM I 6–7). There is a touch of both admiration and horror when Nietzsche writes of 'the ressentiment of beings denied the true reason, that of deed', and of occasions where 'ressentiment itself becomes creative and gives birth to values' (GM I 10).

One might think that perhaps Nietzsche's objection to ressentiment and repression was more on a quasi-moral ground: that one overtaken by ressentiment and repression typically has a false view both of the object of ressentiment and of one's own nature. For instance, Nietzsche claims that repressed resentful slaves imagine, incorrectly, that their masters are free to act other than in a masterly (i.e., slave oppressive) way and that they, the slaves, have freely chosen to be meek, mild, turn the other cheek, etc., 'weakness is to be lied into a merit' (GM I 14). Now for those who, working under the influence of Christian morality, place a categorical value on truth, the objection that repression and ressentiment typically lead their bearers to a false view of themselves and others, counts as probative. In the *Genealogy of Morality* Nietzsche paints a story of repressed slaves becoming creatures of resentment involved in massive self-deception knowing that his readers, as inheritors of Christian morality, will naturally take the charge of self-deception as a severe criticism. But what about those, allegedly such as Nietzsche himself, who strive for an *arche* beyond Christian morality? The demand to be free of self-deception, the demand for truth at any price, does not sit well with Nietzsche's repeated claims that lies and myths are a necessary part of life. Nietzsche repeatedly writes of the need for self-deception (BGE 2, 4) and even claims that ignorance of one's deeper drives and motivations can often be healthy phenomenon (EH, Why I am So Clever 9). Christian morality treats truth as a categorical and ultimate value, that is, as a final value that is higher than all other values. For that reason it espouses a total rejection of all that is opposed to truth, including falsity, myth, and self-deception. While Nietzsche does indeed value truth, he sees it as one of many competing values. Health for Nietzsche involves a certain balancing of these values; taking truth as a value to which everything must be subordinated is for Nietzsche a form of pathology. Nietzsche recognizes that the health of a certain individual or indeed of a certain culture is sometimes built upon certain lies, myths, and self-deception. A Nietzsche who valorized truth above all else would be subject to Nietzsche's own trenchant critique of the ascetic ideal and its fetishization of truth (GM III, *passim*).

It seems reasonable to conclude that Nietzsche does not prefer sublimation over repression because the former facilitates a reduction of misery, and an easing of psychic tension, or because repression leads to false beliefs or even self-deception.

One strand of Nietzsche's arguments against repression is that it prevents a certain kind of unity that he seems to value.[28] It is in this deeper sense that through the effects of repression we become 'strangers to ourselves' (GM, Prologue, 1); we do not simply lack knowledge of ourselves but more importantly we have become estranged from parts of ourselves. Through repression parts of our selves are split-off, not simply in the sense that we have no conscious access to them, but in the sense that we contain within us hidden affects and drives. These are separate movers that are not part of any integrated

whole. Taken to the extreme, this notion of being strangers to ourselves actually threatens the notion of a unified self. That is to say, we have strangers *within* ourselves, so that, in fact, our self is no genuine self. We are nothing more than a jumble of different voices/drives having no overall unity.[29] But what is the basis for valorizing unity? Is it merely an inheritance from the German romantics? Gemes (2006a) argues that Nietzsche values unity because he thinks that it is a necessary condition for freedom and agency. But then why value freedom and agency? Are these some of Nietzsche's ultimate, categorical values?[30]

Alternatively, or additionally, one might argue that Nietzsche, following the romantics, believed that unity is the precondition for great creativity and high culture; these being (some of) Nietzsche's highest values. But Nietzsche himself, along with his story of the repressed, resentful slaves/priests, and his story of Socrates, who create whole new moral universes, seems to run counter to this claim. One suspects that Harold Bloom's thesis that strong poets, and, more generally, other creative types, have a sort of resentment of their predecessors, a certain 'anxiety of influence', which spurs them to their greatest creative deeds, would not at all be uncongenial to Nietzsche. One might argue that in many such cases great creativity comes with the price of a certain denigration of life—for Nietzsche the Christian and Socratic denigration of the non-rational drives is simply a denigration of life. So even allowing that both repression and ressentiment can be spurs to great creativity, Nietzsche would prefer a form of creativity rooted in sublimation as it is more easily wedded to a genuinely life-affirming spirit.

Here is my final suggestion: One of Nietzsche's fundamental grounds for preferring sublimation to repression is ultimately aesthetic. According to Nietzsche repressed persons have parts of themselves that they cannot accept; thus fundamentally ressentiment is really directed at those very parts of themselves. What the repressed person of ressentiment really hates is those parts of himself that his weakness, his inability to act, has led him to repudiate and deny. This self hatred is ultimately disfiguring and ugly. It is intriguing that Nietzsche suggests that his fundamental objection to Socrates was that he was ugly. For Nietzsche, Socrates is a monster of reason who has disfigured himself literally and metaphorically by suppressing all that is not reason in himself:

> In origin, Socrates belonged to the lowest class: Socrates was plebs. We know, we can still see for ourselves, how ugly he was. But ugliness, in itself an objection, is among the Greeks almost a refutation. Was Socrates a Greek at all? Ugliness is often enough the expression of a development that has been crossed, *thwarted* by crossing. Or it appears as declining development. The anthropologists among the criminologists tell us that the typical criminal is ugly: *monstrum in fronte, monstrum in animo* (A, 'The Problem of Socrates', 3).[31]

In the *Gay Science* (290) Nietzsche extols a certain kind of self-satisfaction as the 'one needful thing'. In doing so he raises a distinctly aesthetic objection to that human being

who 'is dissatisfied with himself and is constantly ready for revenge', namely, that we, his potential victims, have to 'endure his ugly sight'.

It is doubtful that Nietzsche had some single ultimate categorical (i.e. end in itself) value. More likely he had a variety of categorical values. Among other things, he valued, as ends in themselves, great individuals, high culture, affirmation of life, expression of power, and various aesthetic values (for instance, beauty). Indeed many of these he valued both categorically and hypothetically (for the things they lead to). For instance, great individuals are valued both as ends in themselves and as a means towards the development of high culture. His strong preference for sublimation over repression is perhaps best understood as embodying his belief that the former provides a more likely path than the latter to realizing these values.[32]

References

Adler, A. (1928). *Über den nervösen Charakter: Grundzüge einer vergleichenden Individual-Psychologie und Psychotherapie*. Munich: Verlag von J. F. Bergmann.

Adler, N. (1986). 'Sublimation and Addiction', *Psychoanalytic Psychology 3*: 187–191.

Arlow, J. (1955). 'Sublimation', *Journal of the American Psychoanalytic Association 3*: 515–527.

Assoun, P. (2000). *Freud and Nietzsche*. London: Continuum.

Bergler, E. (1945). 'On a Five-Layer Structure in Sublimation', *The Psychoanalytic Quarterly 1*: 76–97.

Bos, J. (1992). 'On the Origin of the Id', *International Review of Psycho-Analysis 19*: 433–443.

Chapman, A. and Chapman-Satana, M. (1995). 'The Influence of Nietzsche on Freud's Ideas', *British Journal of Psychiatry 166*: 251–253.

Deri, F. (1939). 'On Sublimation', *The Psychoanalytic Quarterly*, *8*: 325–334

Fenichel, O. (1945). *The Psychoanalytic Theory of Neurosis*. New York: Norton.

Freud, S. (1893–5). *Studies on Hysteria*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 2. London: Hogarth Press.

Freud, S. (1897). 'Letter 61'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 1 (pp. 247–248). London: Hogarth Press.

Freud, S. (1905a). 'Three Essays on Sexuality'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 7 (pp. 123–243). London: Hogarth Press.

Freud, S. (1905b). 'Fragment of an Analysis of a Case of Hysteria' In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 7 (pp. 1–122). London: Hogarth Press.

Freud, S. (1910a). 'Leonardo Da Vinci and a Memory of his Childhood'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 11 (pp. 57–137). London: Hogarth Press.

Freud, S. (1910b). 'Five Lectures on Psychoanalysis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 11 (pp. 3–55). London: Hogarth Press.

Freud, S. (1914). 'On Narcissism'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 67–102). London: Hogarth Press.

Freud, S. (1915). 'Repression'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 141–158). London: Hogarth Press.

Freud, S. (1920). 'Beyond the Pleasure Principle'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 18 (pp. 1–64). London: Hogarth Press.

Freud, S. (1923). 'The Ego and the Id'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 19 (pp. 1–66). London: Hogarth Press.

Freud, S. (1930). 'Civilization and its discontents'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 21 (pp. 57–146). London: Hogarth Press.

Freud, S. (1933). '*New Introductory Lectures on Psychoanalysis*'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 22 (pp. 1–193). London: Hogarth Press.

Gardner, S. (1993). *Irrationality and the Philosophy of the Psychoanalysis*, Cambridge: Cambridge University Press.

Gemes, K. (2006a). 'Nietzsche on the Sovereign Individual, Free Will and Autonomy', *Aristotelian Society* Suppl. *80*: 321–339.

Gemes, K. (2006b). '"We Remain of Necessity Stranger to Ourselves": The Key Message of Nietzsche's Genealogy'. In C. D. Acampora (ed.), *Nietzsche's* On the Genealogy of Morals: *Critical Essays* (pp. 191–208). Lanham, MD: Rowman & Littlefield.

Gemes, K. (2008). 'Nihilism and the Affirmation of Life: A Review of and Dialog with Bernard Reginster', *European Journal of Philosophy 16*: 459–466.

Gemes, K. (2009). 'Freud and Nietzsche on Sublimation', *Journal of Nietzsche Studies 38*: 38–59.

Glover, E. (1931). 'Sublimation, Substitution and Social Anxiety'. *International Journal of Psycho-Analysis 12*: 263–297.

Golomb, J. (1989). *Nietzsche's Enticing Psychology of Power*. Jerusalem: The Magnes Press.

Golomb, J., Santaniello, W., and Lehrer, R. (eds) (1999). *Nietzsche and Depth Psychology*. Albany, NY: State University of New York Press.

Hacker, F. (1972). 'Sublimation Revisited', *International Journal of Psycho-Analysis 53*: 219–223.

Hale, N. (ed.) (1971). *James Jackson Putnam and Psychoanalysis*. Cambridge, MA: Harvard University Press.

Hartmann, H. (1948). 'Comments on the Psychoanalytic Theory of Instinctual Drives', *The Psychoanalytic Quarterly 17*: 368–388.

Hartmann, H. (1955). 'Notes on the Theory of Sublimation', *Psychoanalytic Study of the Child 10*: 9–29.

Heidegger, M. (1979). *Nietzsche*, Vol. I: *The Will to Power as Art*, trans. D. F. Krell. New York: Harper and Row.

Huddleston, A. (2014). 'Nietzsche's Meta-axiology: Against the Skeptical Readings', *British Journal for the History of Philosophy 22*: 322–343.

Kaplan, D. (1993). 'What is Sublimated in Sublimation', *Journal of the American Psychoanalytic Association 41*: 549–570.

Kaufmann, W. (1980). *Nietzsche: Philosopher, Psychologist, Antichrist*, 4th ed. Princeton: Princeton University Press.

Kaywin, L. (1977). 'Problems of Sublimation', *Journal of the American Psychoanalytic Association 25*: 363–370.

Klein, M. (1923). 'The Development of a Child', *International Journal of Psycho-Analysis 4*: 419–474.

Klein, M. (1926). 'Infant Analysis', *International Journal of Psycho-Analysis 7*: 31–63.

Klein, M. (1930). 'The Importance of Symbol-Formation in the Development of the Ego', *International Journal of Psycho-Analysis 11*: 24–39.

Kris, E. (1955). 'Neutralization and Sublimation—Observations on Young Children', *Psychoanalytic Study of the Child 10*: 30–46.

Lacan, J. (1992). *The Seminar of Jacques Lacan Book VII: The Ethics of Psychoanalysis 1959–60*, ed. J. Miller. New York: Norton.

Laplanche, J. and Pontalis, J. (1973). *The Language of Psychoanalysis*. New York: Norton.

Lehrer, R. (1995). *Nietzsche's Presence in Freud's Life and Thought: On the Origins of a Psychology of Dynamic Unconscious Mental Functioning*. Albany, NY: State University of New York Press.

Loewald, H. (1962). 'Internalisation, Separation, Mourning, and the Superego', *The Psychoanalytic Quarterly 31*: 483–504.

Loewald, H. (1973). 'On Internalization', *International Journal of Psycho-Analysis 54*: 9–17.

Loewald, H. (1988). *Sublimation*: *Inquiries into Theoretical Psychoanalysis*. New Haven, CT: Yale University Press.

May, S. (1999). *Nietzsche's Ethics and his War on 'Morality'*. Oxford: Clarendon Press.

Muller, J. (1999). 'Modes and Functions of Sublimation', *Annual of Psychoanalysis 26*: 103–125.

Nietzsche, F. (1966). (BGE) *Beyond Good and Evil*, trans. W. Kaufmann. New York: Vintage.

Nietzsche, F. (1969). (EC) *Ecce Homo*. In *On the Genealogy of Morals and Ecce Homo*, trans W. Kaufmann and R. J. Hollingdale. New York: Vintage.

Nietzsche, F. (1969). (GM) *On the Genealogy of Morals*. In *On the Genealogy of Morals and Ecce Homo*, trans. W. Kaufmann and R. J. Hollingdale. New York: Vintage.

Nietzsche, F. (1975). (TSZ) *Thus Spoke Zarathustra*, trans. R. J. Hollingdale. Middlesex: Penguin.

Nietzsche, F. (1983). (UM) *Untimely Meditations*, trans. R. J. Hollingdale. Cambridge: Cambridge University Press.

Nietzsche, F. (1986). (UM) *Human, All Too Human*, trans. R. J. Hollingdale. Cambridge: Cambridge University Press.

Nietzsche, F. (1988). (KSA) *Sämtliche Werke: Kritische Studienausgabe in 15 Einzelländen*, eds G. Colli and M. Montinari. Berlin: de Gruyter.

Nietzsche, F. (1997). (D) *Daybreak*, eds M. Clark and B. Leiter, trans. R. J. Hollingdale. Cambridge: Cambridge University Press.

Reginster, B. (1997). 'Nietzsche on Ressentiment and Valuation', *Philosophy and Phenomenological Research 57*: 281–305.

Richardson, J. (1996). *Nietzsche's System*. Oxford: Oxford University Press.

Roheim, G. (1943). 'Sublimation', *The Psychoanalytic Quarterly 12*: 338–352.

Sandler, J. and Joffe, G. (1966). 'On Skill and Sublimation', *Journal of the American Psychoanalytic Association 14*: 335–355.

Sartre, J. (1958). *Being and Nothingness*. London: Hazel Barnes.

Schacht, R. (1983). *Nietzsche*. London: Routledge & Kegan Paul.

Segal, H. (1952). 'A Psychoanalytical Approach to Aesthetics', *International Journal of Psycho-Analysis 33*: 196–207.

Segal, H. (1957). 'Notes on Symbol Formation', *International Journal of Psycho-Analysis 38*: 391–397.

Segal, H. (1991). *Dream, Phantasy and Art*. London: Routledge.

Sharpe, E. (1930). 'Certain Aspects of Sublimation and Delusion', *International Journal of Psycho-Analysis 11*: 12–23.

Volney, G. (1992). *Freud on Sublimation—Reconsiderations*. Albany, NY: State University of New York Press.

Notes:

(1) This chapter is a development of Gemes (2009).

(2) Freud often distinguishes between the aim and object of a drive. On one version of Freud's aim/object contrast the aim of a drive is simply satisfaction, or put another way, its aim is to discharge its associated quanta of energy; and the object of a drive is that characteristic activity or things towards which that energy is directed. However, Freud often uses the aim/object contrast to differentiate the characteristic activity of a drive (the aim) from the particular things that drive focuses on at different times (the objects). In this chapter the term 'aim' is typically used in the sense of the active direction of a drive. Thanks are due to Sebastian Gardner for bringing my attention to this point.

(3.) The term Freud invariably used is '*Trieb*' which the translators of the Standard Edition have unfortunately rendered as 'instinct' rather than 'drive'.

(4.) Bergler (1945: 77) similarly notes:

At the time Freud formulated his views on the subject (of sublimation), only the repressed phallic and pre-genital wishes and their contributories were considered to be part of the id. Later, Freud put at least equal stress on repressed aggressive trends. However, no revision of the problem of sublimation was undertaken on the basis of this inclusion of aggressive trends.

(5.) Cf., for instance, Glover (1931: 266–7); Hartmann (1955: 10); Arlow (1955: 515).

(6.) The editors of the Standard Edition hypothesize that among Freud's lost metapsychological papers from 1915–17 there was an essay dealing explicitly with sublimation.

(7.) For instance, as late as his *New Introductory Lectures on Psychoanalysis* from 1933 Freud writes of psychoanalysis as a depth psychology that 'is quite unfit to construct a *Weltanshauung* of its own; it must accept the scientific one', and later in the same work writes 'science is content to investigate and to establish facts' (Freud 1933: 158, 162). While it may be a fact that in society X walking only on the cracks in the pavement is considered not socially valuable, the judgement that it not valuable and hence is a neurotic symptom rather than a sublimation involves an endorsement of the norms of that society and hence is beyond the purview of the mere establishing of facts. Alternatively one is left with the difficult relativist position that all that one can say as a statement of fact is that relative to the norms of society X it is sublimation and not a neurotic symptom. There is no evidence the Freud endorsed this kind of relativism, even if many post-Freudians would be happy to endorse it.

(8.) In *Beyond the Pleasure Principle* Freud gives perhaps his strongest formulation of the claim that sublimations do not involve the removal of repressions: 'The repressed instinct never ceases to strive for complete satisfaction … No substitute or reaction formations and no sublimations will suffice to remove the repressed instinct's persisting tension' (Freud 1920: 42).

(9.) The first known references by Freud to sublimation occur in letters and drafts of letters to Fliess dating from 1897 (Freud 1897: 247). The first reference to sublimation in works published by Freud occurs in the Dora case (Freud 1905b: 50) and in the *Three Essays on Sexuality*, both published in 1905.

(10.) There is a good deal of literature devoted to expounding the relationship between Freud and Nietzsche. Unfortunately, much of it concerns the question of to what degree Nietzsche anticipated Freud and to what degree Freud did or did not cover up influences from Nietzsche. By and large, literature governed by these concerns tends to do little to illuminate the work of either Freud or Nietzsche. Perhaps the best, or, at least, most comprehensive, book of this questionable genre is Lehrer (1995). Less helpful is Assoun (2000). A most succinct example of this genre is A. H. Chapman and Mirian Chapman-Santana (1995).

(11.) Schopenhauer was more prone to speak of wills (*Willen*) rather than drives (*Triebe*). How much this is merely a terminological difference is a difficult question to answer. For Schopenhauer, beyond individual time-located willings, there is, notoriously, the transcendental notion of *Wille*, for which Nietzsche had no sympathy. However, Freud's notion of Eros may, arguably, be seen as a return to a more transcendental picture.

(12.) Note Nietzsche references are given not by page number, save for references to the KSA, but rather by section or aphorism number. For instance 'GM II 16' refers to section 16 of the second essay of Nietzsche's *The Genealogy of Morality*.

(13.) While Nietzsche never takes on anything approximating Freud's topological model of the id, ego, and superego (cf., for instance, *The Ego and the Id*, Freud 1923), Nietzsche countenances different modes of suppression and repression that have echoes in Freud's topological model. For instance, in GM I Nietzsche tells the story of slaves who initially suppress their drives principally because their reality principle tells them they are too weak to act on those drives in the face of the masters' oppression; later in GM Nietzsche tells how that suppression takes on a more moralized form leading to full-blown repression. Indeed, in sections 19–22 of GM II, Nietzsche tells how it is in reaction to our debts to our fathers, forefathers, and God that we develop what he calls a guilty conscience. The first kind of suppression is somewhat akin to suppression seated in the ego and the second kind is akin to repression seated in the superego (itself formed, according to Freud, through the internalization of god-like authoritarian father figures).

(14.) Reginster (1997) also argues that a kind of splitting of self is integral to *ressentiment*, claiming that 'ressentiment corrupts or dis-integrates the self' (301).

(15.) More generally Richardson (e.g. 1996: 25) gives an account of Nietzsche on sublimation that anticipates that given here. It is quite possible that this chapter's original inspiration comes from the excellent Richardson (1996) since, on reflection, that seems to be true of so much of the current author's work. Another source may have been Simon May (1999) who persuasively argues that sublimation, power, and form creation are Nietzsche's key criteria for value. For an amusing, but much more important, case of such 'anxiety of influence' I direct the reader to Bos (1992) which expounds the convoluted relationship between Nietzsche and Freud on the concept of the 'It' or 'Id' as mediated by the curious figure of George Groddeck, author of the somewhat notorious *Book of the It*.

(16.) It's worth noting that in section 205 from *Beyond Good and Evil*, quoted earlier, after describing those 'humans beings of late cultures' who can contain a multitude of drives fighting each other as 'weaker human beings', Nietzsche considers the case where:

a real mastery and subtlety in waging war against oneself, in other words, self control self-outwitting, has been inherited or cultivated then those magical, incomprehensible, and unfathomable ones arise … whose most beautiful expression is found … among artists perhaps in Leonardo da Vinci.

(17.) As Dina Emundts has pointed out to me, and as acknowledged earlier, such stifling, which can be short of full stifling, does not mean the weaker drive need find no form of expression. For instance, in Schreber's case his repressed homosexual drive found outlet in his paranoid fantasies and his transvestism. In the Christian slave, according to Nietzsche, his repression of his desire to have what the masters have is expressed in his proclaimed repudiation of the masters' attainments. The point is that in repression, as opposed to sublimation, those expressions typically take on the logic of opposites—Schreber's desire to be a female is expressed as his being forced by God to be a female against his will, the Christian slave's desire to have power over his masters is disguised as an alleged love of his enemies and his renunciation of the desire for power. So where repressed desires find expression they do so in a way that cannot be integrated into a coherent whole, they represent disintegration not an integration of the subject.

(18.) Note, it is not being claimed here that Nietzsche uses the term 'sublimation' with this exact meaning. While Nietzsche does occasionally use the term '*Sublimieriung*' and cognates in this writings (for instance, GM II 7) he never gives a thematized account of sublimation. The account of sublimation as unification given here is described as a Nietzschean account in that it is in line with many of his uses of that term and, more importantly, it is based on one of Nietzsche's central ideals of health. Even more importantly, this account serves to underwrite the very distinctions between repression and sublimation and between pathological symptoms and sublimations that Freud's account of sublimation fails to underwrite. This reading allows that in certain passages, for instance KSA 12: 254, Nietzsche uses the term '*Sublimieriung*' and cognates in ways more akin to Freud's actual usage and in ways that do not imply the notion of a united self. Schacht (1983) claims that in fact Nietzsche tends to use the term 'spiritualization (*Vergeistigung*) in the sense of sublimation. He further contrasts this to Nietzsche's term 'internalization' (*Verinnerlichung*). This accords with argument presented in the next section that Nietzsche's use of the term 'internalization' is better construed as repression rather than sublimation.

(19.) Similarly, Nietzsche sees the achievement of free will as something open to a limited few. For more on both these themes see Gemes (2006a).

(20.) The French term *Ressentiment* is expressly used by Nietzsche, most notably in GM, essay I.

(21.) For more on this see Gemes (2008) where affirming drives is not explicated in terms of taking a certain cognitive stance, endorsing some positive proposition about one's drives, but rather as fully expressing one's drives.

(22.) One needs to be a little careful here in avoiding the suggestion that Nietzsche favours some static permanent hierarchy among the drives. In fact, Nietzsche often emphasizes the need for a kind of agonal struggle between the drives. Agonal struggle is on the lines of a contest that develops, brings out, the best of the participants, rather than a struggle that leads to their evisceration. While I do not believe there is genuine conflict between the Nietzsche ideal of a unified self and the Nietzschean ideal of a self

engaged in agonal struggle, limitations of space prevent me from developing this point here. Suffice it for now to say that even a master drive, in order to fully develop itself, needs the conflict of robust challenges.

(23.) Kaufmann (1980) similarly lays great stress on the importance of the notion of sublimation for Nietzsche. In ch. 7, 'Morality and Sublimation', Kaufmann explicitly contrasts repression and sublimation as two methods for dealing with the chaos of impulses. Kaufmann places Nietzsche as a strong advocate of the second of these methods. Indeed Kaufmann, while noting that other modern philosophers including Goethe, Novalis, and Schopenhauer used the notion of sublimation, claims that is as Nietzsche who gave the notion the connotation of the transformation of sexual energy that it carries today (Cf. Kaufmann 1980: 219–20).

(24.) Since Nietzsche says little about the nature of master drives there is a certain amount of vagueness about what counts as a unified self for Nietzsche. With the element of social valuation absent there is the worry that, for instance, a reclusive obsessive stamp collector may count as a unified self. For such a person we may envisage a stamp collecting master drive. The answer here is that Nietzsche as a naturalist believes that as humans we come with a rich panoply of inherited drives—this allows that we may also in the course of acculturation acquire new drives. As a matter of empirical fact the reclusive stamp collector will not be a being who is giving expression to all his drives (for instance, drives to sociability, sexual drives, and aggressive drives).

(25.) In fact, some interpreters, including Heidegger (1979) and Adler (1928), insist that for Nietzsche there is one basic drive or force, namely, will to power, and all more specific drives are just modifications of this drive or force. But this is a difficult matter in Nietzsche interpretation that I can't here enter into.

(26.) Generally Freud emphasizes not unity but incompatibility between the demands of various sub-personal agencies, such as the ego, superego, and the id. This is also true of *The Ego and the Id* where he paints a picture of the ego being ever menaced by the conflicting demands of the reality principle, the id, and the superego. It is in this voice that he there labels Eros not as a uniter but as 'a mischief maker' (Freud 1923: 59).

(27.) I owe this observation about the relationship of internal objects to drives to a private correspondence from Bernard Reginster.

(28.) While, as noted earlier, most of Freud's texts on sublimation emphasize notions such as desexualization and social valuations, there are occasional texts in which he, like Nietzsche, emphasizes the unifying element in sublimations. In *The Ego and the Id* Freud tells us that sublimated energy may 'retain the main purpose of Eros—that of uniting and binding—as it helps towards establishing unity, or a tendency to unity' (Freud 1923: 45). However, more typically Freud labels Eros not as a uniter but as 'a mischief maker' (Freud 1923: 59). For more on this see Gemes (2009).

(29.) For more on this theme of self-estrangement see Gemes (2006b).

(30.) Golomb (1989: 71) similarly argues that for Nietzsche sublimation is to be preferred to repression because it allows for a certain kind of freedom, whereas repression tends to transform its bearers into passive, defensive creatures. Indeed Golomb reasonably concludes that while Nietzsche views the ordinary man as a creature of repression, he takes the extraordinary *Übermensh* to be a product of successful sublimation.

(31.) Nietzsche's relationship to Socrates is complex, involving elements of admiration and repudiation. The admiration involves, among other things, recognition of his success in framing a whole new moral universe and his courage in pursuing and facing death for that vision. But even in passages such as GS 340, where we see some of that admiration, Nietzsche characterizes Socrates as a 'monster' and as one who 'suffered life'.

(32.) There is much in this last paragraph that reflects the enormous influence of my colleague Andrew Huddleston on my thinking about Nietzsche. For more on this point see Huddleston (2014). Thanks are also due to Gudrun von Tevenar, Sebastian Gardner, Bernard Reginster, and to the participants of the St John's College Research Centre's Interdisciplinary Seminars in Psychoanalysis for helpful feedback on this chapter, and to Hanna Segal for discussions of her ideas about sublimation and how they differ from Nietzsche's. Initially I approached the topic of sublimation thinking that I could use Freud to clarify Nietzsche on sublimation. It was Sebastian Gardner who, on hearing a paper in which I gave my Nietzschean account of sublimation, helpfully suggested to me that perhaps it would be more advisable to use Nietzsche to clarify Freud on sublimation.

Ken Gemes

Department of Philosophy, Birkbeck College, University of London

Oxford Handbooks Online

Abstract and Keywords

This section of the Handbook consists of four chapters that focus on the different ways in which twentieth-century philosophers engaged with psychoanalysis. The first chapter examines Maurice Merleau-Ponty's existential-phenomenological reformulation of the psychoanalytic unconscious, with emphasis on his argument that the unconscious is an all-pervasive invisible 'atmosphere' (atmosphère) inexorably surrounding the lived-body and shaping all our emotional experience. The second chapter considers Ludwig Wittgenstein's philosophical critique of psychoanalysis and how he drew on Sigmund Freud's writings to test his new linguistic methods in philosophy. The third chapter describes how the Frankfurt School used Freudian psychoanalysis to bolster its Marxist critique of modern society, citing as an example Theodor W. Adorno, who offered an explanation of how fascist mass movements occurred by drawing on Freud's theory of narcissism. The last chapter discusses the key hermeneutic themes found in Paul Ricoeur's engagement with Freud in his book Freud and Philosophy.

Keywords: psychoanalysis, Maurice Merleau-Ponty, unconscious, Ludwig Wittgenstein, Sigmund Freud, philosophy, Frankfurt School, Theodor W. Adorno, narcissism, Paul Ricoeur

Richard G. T. Gipps and Michael Lacewing

WHILE philosophical interest in the unconscious long predated him, the extent of Freud's influence on the twentieth century was such that further philosophical thinking about the nature and significance of unconscious mental life was by way of engagement with or reaction against his thought. This philosophical engagement with psychoanalysis proved a love–hate relationship, and some of the strongest opposing reactions to psychoanalysis, and especially to its father figure, were held within the same individuals or schools. Philosophers related in various ways to psychoanalysis: i) by reformulating or critiquing the theory of the unconscious; ii) by applauding the resource it offers for social critique, or deploring it for reactionary obstructions to emancipation; iii) by vaunting it as a secular spirituality, or decrying its psychological straitjacketing; iv) by emulating its

Subscriber: Cambridge University Library; date: 17 March 2019

methods and aims transposed into the philosophical context; v) by criticizing its failings to measure up to extant paradigms of knowledge, or by drawing on its distinctive insights into human nature to further philosophical projects; vi) by finding within it either a unifying understanding, or a misleading conflation, of matters biological and matters meaningful. The following compressed overview provides a sense of some major lines of development.

i) Classic examples of twentieth-century philosophical reformulations of the unconscious include Maurice Merleau-Ponty's (1970) phenomenological understanding of it as an all-pervasive invisible 'atmosphere' (atmosphère) inexorably surrounding the lived body and shaping all our emotional experience; Martin Heidegger's (2001) ontological refashioning of it in terms of the distinctive form of human existence (Dasein) and the 'clearing' (*die Lichtung*) within which anything shows up for us in experience; Jean-Paul Sartre's (1943/2003) replacing it with bad faith (*mauvaise fois*); and Jacques Lacan's (1966/2002) articulation of it in terms provided by structuralist linguistics.

ii) Philosophers vaunted and decried psychoanalysis as a model for both social and individual emancipation. At the social level the major impetus came from the Marxist philosophy of the Frankfurt School. For example Marcuse (1955) found Freud's focus on the liberation of libidinal life a valuable resource to challenge what is stultifying in the social order, while Fromm (1980) saw Freud's focus on libidinal drives as distracting from the social determinants of psychological illness.

iii) A psychoanalytic ethic of knowing thyself is a famously secular one (see Gipps and Lacewing, 'Introduction: Know Thyself', this volume). Yet twentieth-century philosophers have differed in their assessments of the viability of psychoanalysis as a philosophy for life. Herbart Fingarette (1963), for example, saw in it a profound means of spiritual self-transformation. By contrast Philip Rieff (1966) doubted the moral depths of the 'psychological man' who lives in an 'era of the therapeutic', an era which replaces an ethics of passionate commitment to ideals and social purposes with one of inner, individual, rational, compromise-making self-consciousness. Michel Foucault (1965) similarly considered psychoanalysis as contributing to the creation of its own object of study—a pathologically self-preoccupied subject whose natural outer-directed intentionality is subverted by the inward focus of the psychoanalytic gaze.

iv) Psychoanalysis as a therapy aims not to teach general truths about human life, but to make conscious what an individual has repressed and thereby transcend the obstacles to growth. As a model for philosophical practice this has been most strongly developed in France. Thus Gaston Bachelard (1936/1964) drew on a psychoanalytic model to interrogate and unearth such 'epistemic obstacles', and 'blindly accepted, unquestioned convictions' with an affective rather than rational nature, as inform and constrain scientific theorizing. Jacques Derrida (1998), who often engaged directly with Freud and Lacan in his writing, also developed a 'deconstructive' approach to reading texts which aims to unearth what is repressed yet quietly crucial within an author's system of thought. Luce Irigaray (1985) deploys

a similar method; she pays particular attention to the elision of woman within the dominant forms of thought offered by both psychoanalysis and philosophy. Michel Fouca (1965) analysed reason's duplicitous disavowal of the resources of madness, and sometimes saw psychoanalysis as perpetrator, and sometimes as sensitive recoverer of unreason. In a different spirit, both Ludwig Wittgenstein (2005) and Friederich Waismann (1968) analogized between philosophical practice and psychoanalysis. According to Gordon Baker's (2006) reading, the salient points of comparison involve the non-generalizable, patient-specific nature of a philosophical therapy which is designed not to undo false beliefs but rather to relieve mental cramps which have their roots not in the intellect but in the will.

(p. 135) **v)** The philosophical critique of psychoanalytic science takes various forms. On the one hand we find those—Alasdair MacIntyre (1958) and Adolf Grunbaum (1984), for example—who find psychoanalysis wanting according to standards (e.g. nomological conceptions of causal explanation) invoked, at least for a while, by twentieth-century philosophy of science. On the other we meet with critics—Deleuze and Guattari (1987), for example—who take Freud's psychoanalysis to task for so persistently reading the data through the theory they are supposed to evidence that other ways of reading them become unavailable. And then there are those—Karl Popper (1963), Ludwig Wittgenstein (1966), and Frank Cioffi (1998), for example— who find psychoanalysis variously insightful yet also severally tendentious, especially as it tacitly but effectively renders its own pronouncements immune from meaningful challenge. Contrasting with these critical projects are the uses of psychoanalytic ideas by philosophers to further their philosophical anthropologies. Leaving aside those who themselves appear in the Handbook, the most prominent of these friends of psychoanalysis is perhaps Richard Wollheim (1984) whose writings on personal identity draw deeply on the Freudian and Kleinian theory of identification, the superego, internal objects, and projective identification. Also worthy of mention here is Ilham Dilman's trilogy (1983, 1984, 1988) on Freud and the mind, human nature, insight, and therapeutic change.

vi) Twentieth-century philosophy was preoccupied by the question of naturalism—i.e. with the place of meaning in a natural world. Some philosophers—Susanne Langer (1942) and Paul Ricœur (1970), for example—found in psychoanalysis a happy or unhappy union of man's symbolic and animal natures. Others—such as Ludwig Wittgenstein (1966) and Jurgen Habermas (1972)—complained that psychoanalysis 'conflated reasons and causes' or suffered a 'scientistic self-misunderstanding' when it saw its methods as scientific rather than interpretive, and its insights as belonging to natural science rather than to the humanities.

In this section of the Handbook we offer four exemplary critical and exegetical chapters considering particular twentieth-century philosophical engagements. James Phillips considers Merleau-Ponty's existential-phenomenological retheorization of the psychoanalytic unconscious. Merleau-Ponty's thought is particularly important as it is located at the intersection of two of the most important strands of twentieth-century thinking about human nature—namely Heidegger's existential phenomenology and

Freud's psychoanalysis—and its significance for understanding unconscious life is only now being appreciated (see also Fuchs, this volume). Donald Levy critically examines Wittgenstein's philosophical critique of psychoanalysis. Wittgenstein offered us no sustained treatment of psychoanalysis, but was, like Freud, a product of early twentieth-century Vienna, with a sister treated by Freud; he styled himself a 'disciple of Freud', and kept on returning to Freudian themes throughout his life's work. Freud's writings—especially the *Interpretation of Dreams*—provided him with opportunities to test his (p. 136) new linguistic methods in philosophy. Such methods were designed to reveal latent confusions about (for example) the mind, confusions caused by making simplifying assumptions about how the language of the mind works, assumptions which may be overcome by attending to the rich diversity of its essential forms. Martin Jay surveys the Frankfurt School's sundry uses of Freudian psychoanalysis to bolster its Marxist critique of modern society. An important question for philosophy in integrative mode is how to understand the relations between the trans-individual nature of society and social change and the intrapsychic forces at work in individuals. To give here just one example from Jay's discussion, Adorno drew on Freud's theory of narcissism to explain how fascist mass movements occurred—the authoritarian-type personality both impotently submitting to authority while also magically identifying with it and finding an out-group to serve as the locus for all that is denigrated and abjected. Finally, Richard Bernstein provides an explication of the key hermeneutic themes to be found in Ricœur's engagement with Freud in his *Freud and Philosophy*. Ricœur's book remains one of the most substantial single-volume treatments of Freud's thought available in English, and Bernstein provides an introduction to the significant contours of Ricœur's early study of Freud.

This is a small selection from among the many strands of twentieth-century philosophers' engagement with Freud and psychoanalysis. For readers interested in Heidegger's and Sartre's approaches to psychoanalysis, and in the Daseinanalytic approaches of Medard Boss and Ludwig Binswanger, see Askay and Farquhar (2006) and Holzhey-Kunz (2014). For a clinical introduction to Lacan, see Bailly (2009) and for a philosophical exploration, Grigg (2009). Many of the themes canvased here are still live concerns in the engagement between philosophy and psychoanalysis and as such naturally take up their place in the majority of the chapters comprising the rest of this Handbook. Yet whereas twentieth-century engagements between philosophy and psychoanalysis could sometimes be marked by 'either/or's (either reasons or causes, either a science or a humanity, etc.), a salient feature of many of the twenty-first-century engagements represented here is the work they put in to transcend—or at least contend with, without inappropriately resolving—such dichotomies.

References

Askay, R. and Farquhar, J. (2006). *Apprehending the Inaccessible: Freudian Psychoanalysis and Existential Phenomenology*. Evanston, IL: Northwestern University Press.

Bachelard, G. (1936/1964). *The Psychoanalysis of Fire*, trans. A. C. M. Ross. London: Routledge & Kegan Paul.

Bailly, L. (2009). *Lacan: A Beginner's Guide*. Oxford: Oneworld.

Baker, G. (2006). *Wittgenstein's Method: Neglected Aspects*. Oxford: Wiley-Blackwell.

Cioffi, F. (1998). *Freud and the Question of Pseudoscience*. Chicago, IL: Open Court.

Deleuze, G. and Guattari, F. (1987). *A Thousand Plateaus*, trans. B. Massumi. Minneapolis, MN: University of Minnesota Press.

Derrida, J. (1998). *Resistances of Psychoanalysis*, trans. P. Kamuf, P.-A. Brault, and M. Naas. Stanford, CA: Stanford University Press.

(p. 137) Dilman, I. (1983). *Freud and Human Nature*. Oxford: Blackwell.

Dilman, I. (1984). *Freud and the Mind*. Oxford: Blackwell.

Dilman, I. (1988). *Freud, Insight and Change*. Oxford: Blackwell.

Fingarette, H. (1963). *The Self in Transformation*. New York: Harper & Row.

Foucault, M. (1965). *Madness & Civilization*, trans. Richard Howard. New York: Pantheon.

Fromm, E. (1980). *Beyond the Chains of Illusion: My Encounter with Marx & Freud*. London: Sphere Books.

Grigg, R. (2009). *Lacan, Language and Philosophy*. New York: SUNY Press.

Grunbaum, A. (1984). *The Foundations of Psychoanalysis: A Philosophical Critique*. Berkeley, CA: University of California Press.

Habermas, J. (1972). *Knowledge and Human Interests*, trans. Jeremy J. Shapiro. London: Heinemann.

Heidegger, M. (2001). *Zollikon Seminars*, trans. Franz Mayr and Richard Askay. Evanston, IL: Northwestern University Press.

Holzhey-Kunz, A. (2014). *Daseinsanalysis*. London: Free Publishing Ltd.

Irigaray, L. (1985). *Speculum of the Other Woman*, trans. Gillian C. Gill. Ithaca, NY: Cornell University Press.

Lacan, J. (1966/2002). *Écrits: A Selection*, trans. Bruce Fink. New York: W. W. Norton.

Langer, S. (1942). *Philosophy in a New Key: A Study in the Symbolism of Reason, Rite, and Art*. Cambridge, MA: Harvard University Press.

MacIntyre, A. (1958). *The Unconscious*. London: Routledge & Kegan Paul.

Marcuse, H. (1955). *Eros and Civilisation: A Philosophical Inquiry into Freud*. Boston, MA: Beacon Press.

Merleau-Ponty, M. (1970). *Themes from the Lectures at the Collège de France 1952–1960*, trans. J. O'Neill. Evanston, IL: Northwestern University Press.

Popper, K. (1963). *Conjectures and Refutations: The Growth of Scientific Knowledge*. London: Routledge & Kegan Paul.

Ricœur, P. (1970). *Freud and Philosophy: An Essay on Interpretation*. New Haven, CT: Yale University Press.

Rieff, P. (1966). *The Triumph of the Therapeutic: Uses of Faith after Freud*. Chicago, IL: University of Chicago Press.

Sartre, J.-P. (1943/2003). *Being and Nothingness: An Essay in Phenomenological Ontology*, 2nd edn, trans. H. E. Barnes. London: Routledge.

Waismann, F. (1968). *How I See Philosophy*. London: Palgrave Macmillan.

Wittgenstein, L. (1966). *Lectures and Conversations on Aesthetics, Psychology & Religious Belief*. Oxford: Basil Blackwell.

Wittgenstein, L. (2005). *Big Typescript*. Oxford: Blackwell.

Wollheim, R. (1984). *The Thread of Life*. Cambridge, MA: Harvard University Press.

(p. 138)

Richard G.T. Gipps

Dr Richard Gipps is a clinical psychologist in private psychotherapy practice in Oxford, UK, and an associate of the Philosophy Faculty at the University of Oxford. He convenes the Philosophy Special Interest Group of the Institute of Psychoanalysis, the Oxford Interdisciplinary Seminars in Psychoanalysis, and the Making the Unconscious Conscious seminar series. His research interests lie in psychoanalysis, psychosis, existential phenomenology, and Wittgenstein.

Michael Lacewing

Dr Michael Lacewing is a former Vice-Principal Academic and Reader in Philosophy at Heythrop College, London, an Honorary Reader in Research Department of Clinical, Educational and Health Psychology at University College, London, and a teacher of philosophy and theology at Christ's Hospital School, Sussex. He has published widely in philosophy of psychoanalysis, metaethics and moral psychology, alongside writing textbooks for A level philosophy and training in Philosophy for Children (P4C).

Oxford Handbooks Online

Merleau-Ponty and Psychoanalysis ⓐ

James Phillips

The Oxford Handbook of Philosophy and Psychoanalysis
Edited by Richard G.T. Gipps and Michael Lacewing

Abstract and Keywords

This chapter traces the trajectory of Maurice Merleau-Ponty's views on psychoanalysis and the Freudian unconscious. It begins with a discussion of Merleau-Ponty's first reading of Sigmund Freud by way of The Structure of Behavior, his development of a full description of perceptual consciousness in the Phenomenology of Perception, and his interpretation of the Freudian unconscious as the ambiguity of operative intentionality or perceptual consciousness. The chapter goes on to consider Merleau-Ponty's lectures as Professor of Psychology at the Sorbonne from 1949 to 1952, the criticisms hurled against his early work, and his relationship with Jacques Lacan. It also examines themes from Merleau-Ponty's lectures at the Collège de France between 1951 until his death, along with two of his posthumous writings: The Visible and the Invisible, and a Preface to L'Oeuvre de Freud et son Importance pour la Monde, a book on Freud by psychoanalyst Angelo Hesnard.

Keywords: Maurice Merleau-Ponty, psychoanalysis, Freudian unconscious, perceptual consciousness, Jacques Lacan, phenomenology

Introduction

In this chapter on Merleau-Ponty and psychoanalysis our focus is on Merleau-Ponty's treatment of psychoanalysis and the Freudian unconscious.[1] The 'unconscious' deserves attention because, on the one hand it is, as Laplance and Pontalis point out (1973: 474), the single word that more than any other sums up Freud's discovery; and on the other hand, phenomenology, in its intense study of consciousness, was traditionally sceptical of the notion of an unconscious. Merleau-Ponty, who studied psychoanalysis more intensely than any other phenomenologist, took issue with his colleagues and argued for the introduction of psychoanalysis and the unconscious into phenomenological thought. He did not mean the reified unconscious of Freud, which he also rejected, but a philosophically reformulated unconscious. He regarded the unconscious as a profound insight that Freud was unable to express in philosophically suitable language. Caught up as he was in the language of representational thought, Freud could only think of the unconscious in terms of unconscious representations, 'a tribute', as Merleau-Ponty put it in *Themes from the Lectures at the Collège de France* (hereafter TLF), 'paid by Freud to the psychology of his day' (Merleau-Ponty 1970: 130).

Just as psychoanalysis needs phenomenology to reformulate the unconscious, so also does phenomenology need the psychoanalytic unconscious to reach a depth not achieved in phenomenology. Speaking of this mutual need in an interview with Madeleine Chapsal in his later years, Merleau-Ponty asked:

> Does psychoanalysis render man intelligible? Does it allow us to dispense with philosophy? On the contrary, it poses more vigorously than ever a question that cannot be resolved without philosophy: how can man be at once, completely spirit and completely body? The technique of psychoanalysis contributes to resolving this question along with many other inquiries, and philosophy is again at the crossroads

> (Merleau-Ponty 1992: 6).

The Early Period

The Structure of Behavior

The Structure of Behavior (1963/1942, hereafter cited as SB) represents Merleau-Ponty's first reading of Freud. The first sentence of the book strikes a theme that will remain constant throughout Merleau-Ponty's life: 'Our goal is to understand the relations of consciousness and nature: organic, psychological, or even social' (SB: 3). The terms of this discussion will change, but Merleau-Ponty will always remain focused on this central theme: that of understanding the mysterious bond of consciousness and world, of the interior and the exterior, the subjective and the objective. The goal of *The Structure of Behavior* is to examine this theme from the point of view of the observer. Merleau-Ponty argues that behaviour at any level can only be understood in terms of forms, and that these forms belong neither to nature nor to constituting consciousness, but rather to the spontaneous bond of consciousness and nature that he will eventually call perceptual consciousness. He maintains that the human order of perceptual consciousness is radically distinct from the vital order, that of animal behaviour and its milieu, and that in human behaviour the vital is surpassed and integrated into fully human activity.

Freud is introduced into this discussion as an example of an attempt to reduce the human to the vital—to explain human psychopathology by means of causal, biological forces. To account for psychopathology within the structures of the human order, Merleau-Ponty makes a critical distinction between normally and abnormally structured behaviour. He distinguishes between normal behaviour, in which infantile attitudes are surpassed and integrated into adult behaviour, and pathological behaviour, in which the infantile attitudes have been transcended only in appearance. A 'complex' is a segment of such unintegrated behaviour. In this case the adult lives a life of fragmented consciousness. The phenomena which Freud describes in the categories of conflict, repression, and the unconscious may thus be understood as the life of a fragmented consciousness whose split off segments are present and operative at different times.

This analysis generates a first reformulation of the Freudian unconscious. As Merleau-Ponty puts it: '... the pretended unconsciousness of the complex is reduced to the ambivalence of immediate consciousness' (SB: 179)—ambivalence in that conduct may at one moment reflect an integrated, adult organization and at another moment regression to the more primitive, infantile organization of a complex. Merleau-Ponty concludes *The Structure of Behavior* by invoking the notion of perceptual consciousness, in which the polarity of structured and unstructured, fragmented consciousness is replaced by the ambiguity of perceptual consciousness itself. Merleau-Ponty takes up perceptual consciousness in his second and major work, *Phenomenology of Perception*.

Phenomenology of Perception

Merleau-Ponty does not devote a lot of direct attention to psychoanalysis as such in the *Phenomenology of Perception* (1962/1945, hereafter cited as PP). What he does rather is develop a full description of perceptual consciousness, which will serve as the background for the examination of psychoanalysis and the unconscious both in this work and in that of the following several years. Perceptual consciousness will provide a way of reading Freud, and Freud will offer a confirmation of perceptual consciousness.

For Merleau-Ponty, the method of analysis of perceptual consciousness is phenomenology, but not a phenomenology which posits a transcendental consciousness that constitutes the world in a transparent manner, but rather a phenomenology that emphasizes perception as the lived, embodied experience of a human subject that is inextricably bound up with the world. It is important to note that the subject matter of the *Phenomenology of Perception* is not simply perception, as this is usually understood. It is rather the totality of experience accessible at the level of perception. As Spiegelberg writes: 'Merleau-Ponty's phenomenology of perception is primarily an attempt to explore the basic stratum in our experience of the world as it is given prior to all scientific interpretation' (1982: 561–2). Spiegelberg might also have written 'prior to reflection', inasmuch as perceptual consciousness is first and foremost spontaneous and pre-reflective. The pre-reflective relationship of man and world is described in a variety of ways, such as lived body, being-in-the-world, pre-reflective life, operative intentionality, and existence. Assuming a Heideggerian manner in describing the pre-reflective unity of the subject and world, Merleau-Ponty writes:

> This world is inseparable from the subject, but from a subject that is nothing but a projection of the world, and the subject is inseparable from the world, but from a world which it projects itself. The subject is a being-in-the-world, and the world remains subjective since its texture and articulations are indicated by the subject's movement of transcendence (PP: 430).

In describing perceptual consciousness Merleau-Ponty emphasizes its pre-reflective or unreflective quality, its ambiguity. Consciousness is always opaque, ambiguous, always embedded in uncertainty. In describing this obscurity of one's sense of self, he writes: '… the lived is never entirely comprehensible, what I understand never quite tallies with my living experience, in short, I am never quite at one with myself' (PP: 347).

This analysis of ambiguous consciousness is the basis for Merleau-Ponty's rejection of Freud's understanding of the unconscious as a split off, reified system or, as it were, a second, out-of-awareness 'consciousness'. Merleau-Ponty argues that everything Freud assigns to the unconscious can be understood as pre-reflective consciousness. He writes:

The lived is certainly lived by me, nor am I ignorant of the feelings which I repress, and in this sense there is no unconscious. But I can experience more things than I represent to myself, and my being is not reducible to what expressly appears to me concerning myself. That which is lived is merely ambivalent (PP: 179).

To capture Merleau-Ponty's sense of the unconscious as pre-reflective consciousness rather than as a split off second system, a long quote from a later lecture, 'Man and Adversity', is in order.

In order to account for that osmosis between the body's anonymous life and the person's official life which is Freud's great discovery, it was necessary to introduce something *between* the organism and ourselves considered as a sequence of deliberate acts and express understandings. This was Freud's *unconscious*. We have only to follow the transformations of this Protean idea in Freud's work, the diverse ways in which it is used, and the contradictions it involves to be convinced that it is not a fully developed idea, and that, as Freud himself implies in his *Introductory Lectures on Psychoanalysis*, we still have to find the right formulation for what he intended by this provisional designation. At first glance 'the unconscious' evokes the realm of a dynamics of impulses whose results alone would presumably be given to us. And yet the unconscious cannot be a process 'in the third person'; since it is the unconscious which chooses what aspect of us will be admitted to official existence, which avoids the thoughts or situation we are resisting, and which is therefore not *un-knowing* but rather an un-recognized and unformulated knowing that we do not want to assume. In an approximative language, Freud is on the point of discovering what other thinkers have more appropriately named *ambiguous perception*. It is by working in this direction that we shall find a civil status for this consciousness which brushes its objects (eluding them at the moment it is going to designate them, and taking account of them as the blind man takes account of obstacles rather than recognizing them), and knows about them to the extent that it does not know about them, and which subtends our express acts and understandings

(Merleau-Ponty 1964: 229–30).

To illustrate this experience of spontaneous, ambiguous experience that is only later brought to reflection, Merleau-Ponty offers the example of falling in love. In this experience the subject does not start with an awareness that he will later describe as, 'I was falling in love'. Such an awareness only comes later, when reflecting back on the first steps of the relationship. 'The love which worked out its dialectic through me, and of which I have just become aware, was not, from the start, a thing hidden in my unconscious, nor was it an object before my consciousness, but the impulse carrying me toward someone, the transmutation of my thoughts and behaviour' (PP: 381).

Throughout the discussion of perceptual consciousness Merleau-Ponty develops an interpretation of the Freudian unconscious as effectively the ambiguity of operative intentionality or perceptual consciousness. We just noted in his example of falling in love that the early experience was a pre-reflective impulse towards the other. There was no unconscious idea present at the first moments of falling in love. Merleau-Ponty treats all psychoanalytic examples of the unconscious in the same manner. This interpretation of the unconscious is clearly a further development of his interpretation of the unconscious in *The Structure of Behavior*, where he wrote: '… the pretended unconscious of the complex is reduced to the ambivalence of immediate consciousness' (SB: 179).

Moving this analysis to the area of psychopathology, Merleau-Ponty takes up the condition of hysteria. Freud interprets hysteria as an unacceptable impulse or traumatic memory that is repressed into the unconscious and then forces itself into consciousness as a hysterical symptom, the treatment for which is an uncovering of the repressed idea or impulse. Merleau-Ponty rejects this analysis at every level: for its attempt at an empirical, causal analysis of the hysterical symptom, for its dissociation of the mental and the physical in a way that ignores the unity of the lived body, and for its analysis of behaviour in terms of representational thought, as in Freud's treatment as the recovery of the repressed idea or representation. In contrast, Merleau-Ponty treats hysteria as a form of self-deception in which the hysteric is not fully aware of herself as a lived body and takes distance from her bodily states, not quite treating them as herself. That is why, despite the reported pain or paralysis, the hysteric can appear indifferent to her symptoms. With regard to treatment, Merleau-Ponty again disagrees with Freud's treatment by insight (recovering the repressed memory) and proposes that successful treatment will be below the cognitive, insight-oriented level, and will involve a change in the patient's existence through the relationship with the therapist. Merleau-Ponty recognizes that Freud would dismiss this approach as mere suggestion—or quasi-hypnosis.

Merleau-Ponty carries his analysis one step further in addressing the topic of sexuality, the heartland of Freud's treatment of neurosis. As before, the Freudian unconscious, considered as a region of formed mental representations, is rejected in favour of the ambiguous intentionalities of the lived body. Merleau-Ponty begins his analysis by broadening the discussion of sexuality to that of the entire affective life. Affective experience makes dramatically clear the relations of embodied subject and world. In affective experiences with others, we see that relations with others always involve the self as affective and embodied. That said, sexual relations stand out as the most intense form of affective, embodied experience.

In discussing sexuality Merleau-Ponty first rejects the empirical account that sexuality involves sexual arousal plus a stimulating object. He argues that sexual experience, rather than arising from an autonomous arousal, involves perceiving the other as sexually desirable. He argues further that sexual feeling follows the general trend of the individual's existence. Sexuality reflects one's existence, and one's existence is experienced in one's sexuality; e.g. if one is aggressive in everyday life, one will be

aggressive in one's sexual relations. Merleau-Ponty regards this reciprocity between sexuality and existence as the true discovery of psychoanalysis. While he rejects any reduction of human relations to sexuality, he agrees with Freud that sexuality bears a special relation to neurosis. Whatever conflicts one is experiencing in one's life will be acted out with great intensity in one's sexual relations.

Merleau-Ponty concludes his discussion of sexuality by describing sexuality as an 'ambiguous atmosphere' (PP: 169) that always accompanies existence. What this means is that in sexualized conflicts it is never possible to determine how much the motivation is sexual and how much non-sexual. This point follows Merleau-Ponty's theme that human existence is inherently ambiguous.

The Sorbonne Lectures

A few years after the publication of the *Phenomenology of Perception* in 1945, Merleau-Ponty was appointed Professor of Psychology at the Sorbonne from 1949 to 1952.[2] He lectured extensively on child psychology, and this included a study of post-Freudian psychoanalysis. In this context he examined childhood sexuality, emphasizing the way in which sexuality pervades childhood, but without the child identifying it as such. Of these lectures psychoanalyst Jean Laplanche pointed out 'how much we can learn from Merleau-Ponty. A philosopher interested in clinical observation, in the very concrete experiences involving children, and in the observations of an anthropologist! He could teach a lesson to more than one psychoanalyst' (Laplanche 1989: 92).

In the context of this chapter what we can take from these lectures is another way of understanding the unconscious—in the figure/ground terminology of Gestalt psychology. The point is that the boundary between figure and ground is always blurred. Thus, for both the child and the adult (in different ways), any spontaneous act is surrounded by unfocused-upon intentions, some of which, when brought into awareness (turned from ground to figure), may be found to be influencing the person's actions. Ultimately, of course, the unseen ground represents the ambiguity of pre-reflective life. Merleau-Ponty emphasizes finally, and here he is in full accord with psychoanalysis, that bringing the ground into awareness as figure may require the intervention of another person.

In order to know (ourselves), we need a certain distance that we are not able to take by ourselves. It is not a matter of an unconscious that would play tricks on us; the problem of mystification (of a deceptive consciousness) stems from the fact that all consciousness is the privileged consciousness of a 'figure' and tends to forget the ground without which it has no meaning. We do not know the ground although it is lived by us. We are ourselves our own ground. For knowledge to progress, for there to be scientific knowledge of that 'other' (the ground), it is necessary for what was ground to become figure (Merleau-Ponty 1988: 113).

Critique of Early Work

The presentation of sexuality in the *Phenomenology of Perception*—and along with it the interpretation of the unconscious as unreflective life—was subjected to much criticism. Concerning the issue of sexuality, Merleau-Ponty's effort in the *Phenomenology of Perception* not to 'allow sexuality to become lost in existence' (PP: 159) did not prevent the criticism that he had done exactly that. Psychoanalyst André Green, for instance, wrote: 'The *Phenomenology of Perception* reflected an existential position, and the unconscious as well as sexuality were stripped of the specificity Freud had given them to be dissolved in all too general significations, especially close to those of existential analysis' (Green 1964: 1035). Psychoanalyst and philosopher Jean-Bertrand Pontalis wrote similarly of Merleau-Ponty's analyses that 'the meaning of the symptom dissolves itself in a global existential attitude' (Pontalis 1982: 85).

Arguing in the same critical manner, Paul Ricoeur wrote that the discovery and rediscovery of the unconscious takes place in the treatment situation, that the very notion of an unconscious is intimately associated with the phenomena of conflict, repression, and resistance, and can't be understood out of that context. And further that if the unconscious emerges through analytic work on resistance and repression, it presents itself not as pre-reflectively conscious but rather as separate, autonomous, and heterogeneous—as another system, in Freud's terms. Ricoeur summarizes his argument: 'The justification for the topographic differentiation into systems is to be found in praxis; the "remoteness" between the systems and their separation by the "barrier" of repression are the exact pictorial transcriptions of the "work" that provides access to the area of the repressed' (1970: 413).

The conclusion to this line of critique is that the only way to view the 'unconscious' of the *Phenomenology of Perception* is as Freud's preconscious, as a region of experience that is not in consciousness but that is eminently recoverable through reflection. As Ricoeur put it: 'The unconscious of phenomenology is the preconscious of psychoanalysis, that is to say, an unconscious that is descriptive and not yet topographic' (Ricoeur 1970: 392). And in this vein Pontalis pointed to the distinction between the Freudian unconscious and Merleau-Ponty's pre-reflective life as that of 'the other scene' versus 'the other side' (Pontalis 1971: 62), and he made it clear that the Husserlian, phenomenological 'other side' is the psychoanalytic preconscious.

Was this matter thus settled with the conclusion that Merleau-Ponty's pre-reflective consciousness only reaches Freud's preconscious? Not quite. To the Freudian position that while the repressed unconscious is accessible only through the agency of another, the preconscious is accessible to self-reflection, Merleau-Ponty responded, as noted earlier, that dimensions of pre-reflective consciousness (the 'ground' in the terminology of the *Sorbonne Lectures*) may *not* be accessible to self-reflection and may require the distance of another to see them. It is arguable that most of the features of the dynamic unconscious could be found in Merleau-Ponty's approach to pre-reflective life. In fact,

Merleau-Ponty and Psychoanalysis

Belgian philosopher/psychoanalyst Alphonse de Waelhens and French psychoanalyst Georges Lanteri-Laura both argued in that direction—that the phenomena described in terms of the Freudian unconscious require not so much a second system as an opacity in pre-reflective consciousness. De Waelhens (1966) pointed out that the intentional relations between man and world are so greatly varied that no simple distinction between conscious and unconscious experience is possible, and Lanteri-Laura described the unconscious as that dimension of experience not recovered through reflection:

> The unconscious, we can say, is the entire part of non-thetic [undirected] consciousness that man is not able to recover through reflection alone and that he apprehends only through singular noematic [specific] qualities of objects. If man is essentially a movement of transcendence toward that which he is not, he apprehends himself in totality through the noematic qualities of objects and of others, without moreover knowing it specifically, and it is this hiatus that, in *Phenomenology of Perception*, renders possible the very notion of an unconscious
>
> (Lanteri-Laura 1966: 399).

Thus far in this section we have reviewed the critique—and defence—of Merleau-Ponty's view of the unconscious by a number of his contemporaries in psychoanalysis and philosophy. However, the main challenge—and confrontation—came from Jacques Lacan.

Merleau-Ponty and Lacan

The relationship of Merleau-Ponty and Lacan dates back to Alexandre Kojeve's 1933 lectures on Hegel in Paris (Roudinesco 1990: 350), which both attended and which were a significant influence on both of them.[3] Given that Lacan had a serious interest in philosophy and Merleau-Ponty an equal interest in psychoanalysis, it's hardly surprising that they remained in contact over the ensuing years. They clashed, however, over their respective views on the unconscious. While Lacan regarded Merleau-Ponty's interpretation of the unconscious as hopelessly embedded in a philosophy of consciousness, Merleau-Ponty regarded Lacan's view as an idealist deviation in psychoanalysis.

There were three direct encounters during the 1950s, and then implicit encounters following Merleau-Ponty's death in 1961. In the first encounter in January 1955, Merleau-Ponty gave a lecture to the Société française de Psychanalyse entitled 'Philosophy and Psychoanalysis'. We don't have the text of Merleau-Ponty's presentation, but we do have Lacan's discussion of the lecture in his seminar the next day. Lacan challenged Merleau-Ponty's interpretation of the unconscious as a dimension of a unified consciousness. '… for Merleau-Ponty, it's all there, in consciousness' (Lacan 1988/1954–5: 77–8).

The second episode occurred in 1957, when Lacan presented a paper to the Société française de Philosophie entitled 'Psychoanalysis and its Teaching'. Merleau-Ponty attended the lecture and engaged Lacan in a lively discussion following the lecture.

Lacan challenged the philosophers regarding the unconscious: 'In the unconscious, which is less profound than inaccessible to the depths of consciousness, *it (Id) speaks*: a subject in the subject, transcendent to the subject, poses its question to the philosopher since *The Interpretation of Dreams*' (Lacan 1957: 65). The debate that followed the lecture centred on how to understand a parapraxis, taking as an example one from Freud's own personal history, as related by Lacan. While Merleau-Ponty argued that a parapraxis is a failure of language use, Lacan countered that it is a successful use of language in that it allows the unconscious to speak.

A final encounter between Merleau-Ponty and Lacan took place in 1960, some months prior to Merleau-Ponty's death, at psychiatrist Henri Ey's annual colloquium at Bonneval, that year devoted to the unconscious (Ey 1966). The colloquium has been described at length by Elisabeth Roudinesco in her *Jacques Lacan & Co.* (1990). Prominent philosophers, psychiatrists, and psychoanalysts gathered to discuss the status of the Freudian unconscious.

> For Lacan, what was at stake at Bonneval was considerable. It was a question of demonstrating in France, in the teeth of the IPA [International Psychoanalytic Association], that Freudianism as revised and corrected by linguistics had the full status of a science. If philosophy wanted to escape its rut, it would have to interrogate psychoanalysis and admit that the Freudian unconscious placed the certitudes of consciousness in jeopardy
>
> (Roudinesco 1990: 308).

For Lacan the conference was mostly a success. For the philosophers in attendance, however, the result was mixed. Although there was some agreement on the importance of the Freudian unconscious, not all agreed with Lacan's formulation of it. Primary among the detractors was Merleau-Ponty, the most prominent philosopher in attendance.

Because of Merleau-Ponty's unexpected death months after the meeting, he wasn't able to write his own summary of his remarks. That task fell to Pontalis (with the concurrence of Merleau-Ponty's wife, herself a Lacanian psychoanayst), who reported that Merleau-Ponty began with an interpretation of the unconscious as a 'primordial symbolism', and then continued:

> Only, this primordial symbolism, must we not seek it, rather than in language as such—'it makes me uneasy to see the category of language occupy the entire field'— in a certain perceptual articulation, in a relation between the visible and the invisible that M. Merleau-Ponty designates by the name of latency, in the sense that Heidegger gives to this word [*Verborgenheit*], and not in order to specify a being that would conceal itself behind the appearances. Perception, on condition of not conceiving it as an operation, as a mode of representation, but as the double of non-perception, can serve as a model, and even the simple fact of

> seeing: 'to see, this is to have no need to form a thought'. M. Merleau-Ponty recalls that in his view, the opening to being is not linguistic: it is in perception that he sees the birthplace of the word

> (Merleau-Ponty 1966: 143).

Roudinesco relates Lacan's unhappiness with Merleau-Ponty's comments. He took the occasion of the publication of the proceedings in 1966 to falsely accuse Pontalis of transcribing Merleau-Ponty's words incorrectly (Roudinesco 1990: 314).[4]

The Late Period

Themes from the Lectures at the Collège de France

In 1951 Merleau-Ponty assumed a professorship at the Collège de France, a position he would hold until his death. In two of his courses he introduced his further views on the unconscious. Indeed, as Green remarked, his changing view of psychoanalysis and the unconscious was not simply a reflection of his changing philosophy, but may have been one of the fermenting elements in that change (Green 1964: 1030).

In the course on 'Passivity', in a discussion of Sartre's discussion of the dream as pure imagination in which the world is negated, he argues, contra Sartre, for an ambiguous relationship in which a dream is itself a form of consciousness, and waking consciousness itself has an oneiric character. 'Others are present to us in the way that dreams are, the way myths are, and that is enough to question the cleavage between the real and the imaginary' (TLF: 48). Merleau-Ponty is now introducing something new. Instead of describing the unconscious as ambiguous consciousness, he is now talking about oneiric consciousness. Elaborating on this, he writes rather dramatically that Freud's 'most interesting insight' is the notion of a 'primordial symbolism … the idea of a "non-conventional thought" … which is the source of dreams and more generally of the elaboration of our life' (TLF: 49).

In this discussion Merleau-Ponty apparently doesn't recognize that he is now talking about what for Freud is one of the principle features of the unconscious, namely its character of primary process thinking (nor does Merleau-Ponty use that phrase). For Freud there was, in fact, a conflict as to how to think of unconscious thinking: whether as fully developed unconscious thoughts, or as more primitive, primary process thoughts. In Merleau-Ponty's early writings he focused on the first meaning of the unconscious—as split off unconscious thoughts—and his critique of Freud was directed at that Freudian unconscious, replacing it with his notion of ambiguous consciousness. In the course on Passivity, he now shifts to the other meaning of the unconscious—as primary process. The difference with Freud, now, is that Merleau-Ponty doesn't locate oneiric consciousness

solely in the unconscious; he intermingles it with clear, waking consciousness. This means that so-called waking consciousness can be infused with fantasy, imagination, dream-like material—or, in other words, that waking consciousness is also in part oneiric consciousness.

Merleau-Ponty takes the next step in this train of thought in his last Collège de France course, 'Nature and Logos'. He begins with a focus on the body. Now, however, he takes a radical step that will drive all his late thought: namely, he replaces the 'lived body' of the *Phenomenology of Perception* with a new category of *flesh*. Merleau-Ponty's description of *flesh* is certainly difficult to understand, and it may help to appreciate in advance that with this concept he is embarking on a new, monistic ontology. Although Merleau-Ponty always thought of consciousness and world as thoroughly entwined, he now brings them totally together as two aspects of the same flesh.[5]

We see this shift both in the course on 'Nature and Logos' and in a number of Merleau-Ponty's late and posthumous writings. Writing about corporeal consciousness in 'The Philosopher and his Shadow', Merleau-Ponty writes typically:

> Thus I touch myself touching, my body accomplishes 'a sort of reflection'. In it, through it, there is not just the unidirectional relationship of the one who perceives to what he perceives. The relationship is reversed, the touched hand becomes the touching hand, and I am obliged to say that the sense of touch here is diffused in the body—that the body is a 'perceiving thing', a 'subject–object'
>
> (Merleau-Ponty 1964b/1960: 166).
>
> The other person appears through an extension of that copresence [of the two hands], he and I are like organs of one single intercorporeality
>
> (Merleau-Ponty 1964b/1960: 168).

Flesh is the body as reflexive, as seeing–seen, as touching–touched, as the 'idea of corporeality as an entity with two faces or two sides' (TLF: 129). As he describes this new vision of the body in a very late essay, 'Eye and Mind':

> There is a human body when, between the seeing and the seen, between touching and the touched, between one eye and the other, between hand and hand, a blending of some sort takes place—when the spark is lit between sensing and sensible, lighting the fire that will not stop burning until some accident of the body will undo what no accident would have sufficed to do
>
> (Merleau-Ponty 1964b: 163–4).

The opposition of consciousness and nature has thus collapsed into a phenomenology of flesh.

In 'Nature and Logos' Merleau-Ponty now addresses an obvious question. In a philosophy of the flesh, what becomes of the unconscious? It was originally ambiguous consciousness, then in the course on 'Passivity', oneiric consciousness. Now Merleau-Ponty reformulates it again. He says in this course that the idea of 'unconscious representations' has no place in a philosophy of the flesh and that 'the unconscious is sensing itself, since sensing is not the intellectual possession of "what" is felt but a dispossession of ourselves in favor of it, an opening toward that which we do not have to think in order that we may recognize it' (TLF: 130). In the context of a monistic philosophy of the flesh, sensing becomes Merleau-Ponty's word for the deep and 'unconscious' phenomenon of touching–touched or seeing–seen.

In this context, at the end of the course Merleau-Ponty discusses the psychoanalyst, Melanie Klein, for whom phenomena such as introjection and projection are in the early child bodily experiences—experiences he refers to as 'primordial symbolism'. He now distinguishes the unconscious of repressed phenomena as a 'secondary formation', to be distinguished from a 'primordial unconscious' that 'would be a permissive being, the initial yes, the undividedness of sensing' (TLF: 131). He is making contact with what Klein understood as an initial unity of mother and infant in which the infant doesn't grasp their difference. Oneiric or sensing consciousness thus points to an enduring bond and unity with the other in adult life. To quote from the posthumous Preface: 'Consciousness is now the "soul" of Heraclitus, and Being, which is around it rather than in front of it, is an oneiric Being, by definition hidden' (Merleau-Ponty 1982: 70, hereafter cited as Pref.). Referring to this persistence of primitive or primordial level in adult life, Pontalis remarks of Merleau-Ponty's final philosophy, 'nothing is ever "possessed", but nothing is ever "lost"' (Pontalis 1971: 63), and Green contrasts the role of loss in Freud with the undividedness of Being in Merleau-Ponty (Green 1964: 1032).

Posthumous Writings: *The Visible and the Invisible*

In *The Visible and the Invisible* (1968/1964, hereafter cited as VI), Merleau-Ponty invokes psychoanalysis frequently in the further radicalizing of his thought. In reflecting on this volume, especially the Working Notes that are included at the end of the written text, we must bear in mind that the Working Notes are indeed experimental notes; they are mostly incomplete phrases whose meaning can be very obscure and uncertain. They are also written in a random manner, and it is the reader's challenge to order them in a comprehensible way.

In pursuing his vision of a monistic philosophy of the flesh, Merleau-Ponty corrects his own earlier philosophy, telling us that he is not only abandoning ambiguous consciousness for an ontology of the flesh, but also abandoning consciousness itself. He writes that the problems of the *Phenomenology of Perception* are 'due to the fact that I retained the philosophy of "consciousness"' (VI: 183). The results of *Phenomenology of*

Perception need to be brought to 'ontological explication' (VI: 182). They will then be deepened and rectified (VI: 168).

He describes the new ontology as one in which the Being of flesh splits ('dehisces') into self and world, both participating in the same flesh. Merleau-Ponty develops a special language for this phenomenon, speaking of dehiscence, dimensions, hinges, pivots, reversibility, and chiasms. These terms are all efforts to describe the monistic world in which self and other, self and world, become reversible, and in which our conventional language no longer applies. He writes, for example:

> [P]erhaps the self and the non-self are like the obverse and the reverse and … perhaps our own experience *is* this turning around that installs us far indeed from 'ourselves'. In the other, in the things. Like the natural man, we situate ourselves in ourselves *and* in the things, in ourselves *and* in the other, at the point where, by a sort of *chiasm*, we become the others and we become world (VI: 160).

Writing of oneself and the other he writes further that 'our intentional life involves the *Ineinander* of the others in us and of us in them' (VI: 180). The self and other are no longer two separate individuals facing one another. Rather, the relationship of self and other involves a kind of reversibility, likened to the finger of the glove that is turned inside out.

> In reality there is neither me nor the other as positive, positive subjectivities. There are two caverns, two opennesses, two stages where something will take place—and which both belong to the same world, to the stage of being … There is not the For Itself and the For the Other. They are each the other side of the other. This is why they incorporate one another: projection-introjection—There is that line, that frontier surface at some distance before me, where occurs the veering I-Other Other-I (VI: 263).

What now is the role of psychoanalysis and the unconscious in this discussion? Merleau-Ponty writes that there is a mutual encroachment of philosophy and psychoanalysis. Philosophy will correct psychoanalysis and psychoanalysis will correct philosophy. On the one hand, psychoanalytic terminology has infiltrated Merleau-Ponty's new ontology. Polymorphism and promiscuity are no longer attributes of infantile sexuality but of Being itself. We hear of 'the vertical or carnal universe and its polymorphic matrix' (VI: 221), of 'the fact that the world, Being, are polymorphism, mystery and nowise a layer of flat entities of the in itself' (VI: 252), and that 'the in itself-for itself integration takes place not in the absolute consciousness, but in the Being in promiscuity' (VI: 253). In another note Merleau-Ponty charges himself, 'Do a psychoanalysis of Nature: it is the flesh, the mother' (VI: 267).

On the other hand, if psychoanalysis has encroached on philosophy, it is also the case that philosophy encroaches on psychoanalysis. Thus he writes that 'a philosophy of the flesh is the condition without which psychoanalysis remains anthropology' (VI: 267), and finally

he enjoins himself, 'Make not an existential psychoanalysis but an *ontological* psychoanalysis' (VI: 270).

We can only guess at the meaning of some of these obscure phrases. In speaking of the polymorphism and promiscuity of being/flesh, Merleau-Ponty is certainly using psychoanalytic concepts to reflect the complicated state of a flesh that is both man and other, and man and world. When he directs us to do a psychoanalysis of nature, he is presumably telling us again that the psychoanalytic terms can refer to the world as flesh. On the other hand, when he directs us to make psychoanalysis ontological, it is not at all clear what he means—except that he does *not* mean existential analysis.

What then of the unconscious? The abundance of references to the unconscious in the working notes indicates that Merleau-Ponty's new ontology of the flesh will involve the unconscious more than ever before in his thought, albeit in a very different manner. Recall the trajectory of the unconscious in Merleau-Ponty's thought. From the beginning he rejected the reified Freudian unconscious. He understood the unconscious as a fragmented consciousness in *The Structure of Behavior*, as ambiguous perceptual consciousness in the *Phenomenology of Perception*, as unseen ground in the Sorbonne Lectures, and as oneiric consciousness and sensing in the Collège de France courses. In the working notes the unconscious takes on a new role as the structure that makes relationships possible. He writes:

> One always talks of the problem of the 'other', of 'intersubjectivity' etc. … In fact what has to be understood is, beyond the 'persons', the existentials according to which we comprehend them, and which are the sedimented meaning of all our voluntary and involuntary experience. This unconscious is to be sought not at the bottom of ourselves, behind the back of our 'consciousness', but in front of us, as articulations of our field. It is 'unconscious' by the fact that it is not an *object*, but it is that through which objects are possible, it is the constellation wherein our future is read—It is between them as the interval of the trees between the trees, or at their common level (VI: 180).[6]

We follow Renaud Barbaras here as he emphasizes the displacement of the unconscious away from the subject as well as its structuring of all worldly phenomena. 'The unconscious is rather on the side of the world rather than at the center of the psyche; between things as the pivot or structure around which their phenomenality is constituted, it is a synonym for dimension' (Barbaras 1991: 315). As Merleau-Ponty indicated in the preceding quote, he also refers to the unconscious as an 'existential', with obvious reference to Heidegger's use of this term to describe essential features of Being-in-the-World.

Merleau-Ponty develops this train of thought further in the language of visibility and the invisible, the visible world held together by an invisible structure—by 'an invisible inner framework, and the in-visible is the secret counterpart of the visible, it appears only within it, it is the *Nichturpräsentierba*r which is presented to me as such within the world …' (VI: 215). The invisible framework and the existentials are all but indistinguishable.

Merleau-Ponty and Psychoanalysis

The language of reversibility, chiasm, dimensions, levels, pivots, and hinges applies to both of them. Merleau-Ponty describes the existentials as 'the armature of that "invisible world" which, with speech, begins to impregnate all the things we see' (VI: 180). And of course, it is difficult to distinguish either of these concepts from the newly formulated unconscious.

At the end of this discussion of *The Visible and the Invisible*, it is important to point out what appears to be a persistent ambivalence in Merleau-Ponty's thought. In his late writings he has consistently moved in the direction of a monistic ontology. The Working Notes are the most radical expression of that effort. There remains, however, an ambivalence as to how far he wishes to push the monism. Regarding, for instance, the relations with others, does he really want to argue that I and the other person are simply one person, as opposed to two persons who are intertwined in ways we have not recognized before in philosophy? In a statement quoted earlier, he writes that 'In reality there is neither me nor the other as positive, positive subjectivities. There are two caverns, two opennesses, two stages where something will take place—and which both belong to the same world, to the stage of being …' (VI: 265). He is saying both that there are not two positive subjectivities, but that there *are* two opennesses or two caverns. We are left to wonder how he would have worked out this ambivalence over monism and otherness. He was certainly using the Working Notes to experiment with the monist extreme. It is in the Preface that we can see the ambivalence on full display.

Posthumous Writings: Preface to Hesnard

In 1960 Merleau-Ponty accepted an invitation to write a Preface to a book on Freud by psychoanalyst Angelo Hesnard, *L'Oeuvre de Freud et son Importance pour la Monde* (Merleau-Ponty 1982–3). His agreement to write the Preface speaks to his persistent interest in the psychoanalytic enterprise and provides a valuable summary of his thought, as he calls on his early writings about psychoanalysis and attempts to integrate them into his later thought.

What is at first surprising about the Preface is the degree to which Merleau-Ponty clings to his original point of view on psychoanalysis. Much of the Preface might have been written at the time of the *Phenomenology of Perception*.

He begins the Preface by attributing to the reader what was probably his own history of reading Freud:

> Every reader of Freud, I imagine, remembers his first impression: an incredible bias in favor of the least probable interpretations; a maniacal penchant for the sexual; and above all a distortion, through the use of archaic forms, of significations, language, and action in the interest of derisive puns. Then, to the extent that one read, that one related oneself to oneself, and that the years

passed, a sort of evidence for psychoanalysis was inexplicably established and one came to live in peace with the pitiless hermeneutic (Pref: 68–9).

That said, Merleau-Ponty hastens to add that the psychoanalysis one once refused is not the psychoanalysis one now accepts. What is now accepted is psychoanalysis delivered of Freud's empiricism, biologism, and psychologism. Merleau-Ponty agrees with the author of the book that the theory and practice of psychoanalysis lead to a new philosophy different from that espoused by Freud, and which tends to converge with phenomenology.

In their converging efforts, he continues, each may offer the other something valuable. Phenomenology brings to psychoanalysis a new set of categories, allowing it to see 'the Freudian unconscious as an archaic or primordial consciousness, the repressed as a zone of experience that we have not integrated, the body as a sort of natural or innate complex, and communication as a relation between incarnate beings of the sort who are well or badly integrated' (Pref: 67). Psychoanalysis for its part will bring to phenomenology a depth of experience not reached in a philosophy of consciousness. Merleau-Ponty adds further that we owe to Freud the discovery of the remarkable exchange of roles that is part and parcel of human development—that one can at once be father and son, in becoming a father identifying with both his son and his own father. As Merleau-Ponty puts it: 'This prodigious intuition of exchanges—exchange of roles, exchange of the soul and the body, of the imaginary and the real—this universal promiscuity' (Pref: 68).

He concludes the Preface by questioning whether phenomenology and psychoanalysis converge completely, and whether they do this through a phenomenological translation of psychoanalysis. He answers 'no' to both questions:

> The accord of phenomenology and of psychoanalysis should not be understood to consist in phenomenology's saying clearly what psychoanalysis had said obscurely. On the contrary, it is by what phenomenology implies or unveils at its limits—by its *latent content* or its *unconscious*—that it is in consonance with psychoanalysis. Thus the cross validation between the two doctrines is not exactly *on* the subject man; their agreement is precisely in describing man as a work site (*chantier*), in discovering, beyond the truth of immanence, that of the *Ego* and its acts, that of consciousness and its objects, of relations which a consciousness cannot sustain: man's relations to his origins and his relations to his models. Freud points his finger at the *Id* and the *Superego* ... Phenomenology and psychoanalysis are not parallel; much better, they are both aiming toward the same *latency* (Pref: 71).[7]

Much of what we have just read from the Preface would seem to apply to Merleau-Ponty's early work, as if he were still dealing with perceptual consciousness and the reified Freudian unconscious, rather than with an ontology of the flesh. To the extent that Merleau-Ponty moves beyond that in the Preface, we may find it in his concept of latency. (The other words he uses for latency in the Preface are the 'hidden' (reflecting

Subscriber: University of Otago; date: 17 December 2018

Heidegger's *Verborgenheit*) and Husserl's *Tiefenleben*). In Merleau-Ponty's reading, both psychoanalysis and phenomenology have changed in how they understand what is latent or hidden.

Regarding psychoanalysis, in the Preface Merleau-Ponty reviews his critique of Freud's mechanistic language and his concept of a reified unconscious. He rejects Freud's scientist talk of mnemic traces, projections, unconscious representations, etc., and argues that Freud's genius was in discovering in the clinical encounter a 'treasure of experience'—'in his contact with things, his polymorphous perception of work, of acts, of dreams, of their flux and their reflux, of counter-coups, of echoes, of substitutions, of metamorphoses' (Pref: 68). He rejects any talk of a biological phallus and argues that Freud reveals in work 'an imaginary phallus, a symbolic phallus, oneiric and poetic' (Pref: 69).

Regarding phenomenology, Merleau-Ponty argues that with the publication of Husserl's later work phenomenology has moved beyond the 'transparent correlations between acts of thought and objects' that might be taken from the early work. Phenomenology ceases to be a straightforward relation of consciousness to its objects:

> All consciousness is consciousness of something or of the world, but this *something*, this *world*, is no longer, as 'phenomenological positivism' appeared to teach, an object that is what it is, exactly adjusted to acts of consciousness. Consciousness is now the 'soul of Heraclites', and Being, which is around it than in front of it, is a Being of dreams, by definition hidden: Husserl sometimes uses the term 'pre-being' (Pref: 70).

What does Merleau-Ponty mean with these enigmatic statements? These citations suggest that he is using the term 'latency' to understand further, and beyond the understanding of the *Phenomenology of Perception*, what phenomenology and psychoanalysis, each in its own way, tries to reach beyond consciousness.

Recall now that the final position of Merleau-Ponty's late philosophy is that of a dehiscence of Being into subject and object in which the two are both reversible and united in the notion of flesh—a flesh that is both the flesh of the human subject and the flesh of the world. For Merleau-Ponty, philosophy and psychoanalysis will each get to this core realization of an underlying unity in its own way: philosophy/phenomenology through its analysis of perception, and psychoanalysis through its analysis of interpersonal relations.

In the philosophic investigation of perception, latency emerges in the analysis of the perceptual relation of man and world. As Merleau-Ponty writes, 'It is not we who perceive, it is the thing that perceives itself yonder … The world is a *field*, and as such is always open' (VI: 185). The usual perceptual relation that proceeds from man to world is now transcended. Rather than saying, I perceive, we now say, there is perception—or something like, Being perceives itself through me. Such perception is now articulated in terms of hinges, dimensions, encroachment, and levels of being—and also as the invisible

of the visible, the interior of the exterior. What is latent is the ground of Being or flesh that founds the perceptual relation. If the perceptual relation is thought of as consciousness, then Being or flesh is the latent 'unconscious'. The unseen ground is no longer pre-reflective life but Being (or flesh) itself.

Psychoanalysis finds its own unconscious (or latency) not in the investigation of perception, but rather in the interpersonal relationship, in what earlier was called erotic intentionality and later the libidinal body. In the context of a philosophy of Being in dehiscence, however, we no longer hear psychoanalytically of affective, intentional bonds but rather of a 'prodigious intuition of exchanges', a 'universal promiscuity' (Pref: 68). Thus the psychoanalytic unconscious now refers not to an affective attachment to the other but to a unity with the other through a 'fundamental polymorphism' in which the 'I-other relation [is] to be conceived (like the intersexual relations, with its indefinite substitutions) … as complementary roles one of which cannot be occupied without the other being also: masculinity implies femininity, etc' (VI: 220–1). As indicated earlier, this undividedness of self and other is expressed through a primordial symbolism of the body —what Merleau-Ponty also describes as 'symbolic matrices, a language of self to self, systems of equivalences built up by the past, effecting groupings, abbreviations, and distortions in a simple act and which analysis reconstitutes more and more closely' (Pref: 69). What Merleau-Ponty studied in the Sorbonne Lectures as the primitive unity of mother and infant now persists into adult life. The primordial symbolism of this bond is now the psychoanalytic unconscious (or latency).

Conclusion

In tracing the trajectory of Merleau-Ponty's thought we have witnessed a striking transformation of his view of psychoanalysis and the psychoanalytic unconscious, leading to two possible, and somewhat opposing, conclusions. First, while it is clear that he wanted to turn psychoanalysis in the direction of an ontology of the flesh, it is not clear that this transformation will serve psychoanalysis or psychoanalytic therapy well. As a praxis, as a basis for any psychoanalytically oriented treatment, it may be that Merleau-Ponty's early reading of psychoanalysis will prove more useful than the ontological reading of the later works. In the ongoing confrontation of phenomenology and psychoanalysis, the question will remain whether Merleau-Ponty's early vision of ongoing, perceptual consciousness—pre-reflective, uncertain, ambiguous, never at one with itself—will do the work that Freud assigned to the unconscious and thus provide a clinically useful phenomenological revision of Freud's central idea.

But second, it may also be the case that Merleau-Ponty's later reading of psychoanalysis may add something—a further depth—to psychoanalytic practice. This would involve the discovery of a deeper self—other unity beneath surface relations and conflicts—a unity designated as latency and grounded in an ontology of flesh. Such a proposal is consistent with what in fact has been the post-Freudian history of psychoanalysis, with both its changing focus from Oedipal to pre-Oedipal relationships, and its corresponding emphasis on the therapeutic relation as less asymmetric than Freud viewed it—and perhaps, more than less asymmetric, as also subject to the reversibility of self and other described by Merleau-Ponty (see, e.g. Aron 1996; Loewald 1980; Mitchell 1988).

References

Aron, L. (1996). *A Meeting of Minds: Mutuality in Psychoanalysis*. London: The Analytic Press.

Barbaras, R. (1991). *De l'être du Phénomène: Sur l'Ontolgie de Merleau-Ponty*. Paris: Jérôme Millon.

De Waelhens, A. (1966). 'Sur l'Inconscient et la pensée philosophique'. In H. Ey (ed.), *L'Inconscient, VIᵉ Colloque de Bonneval* (pp. 371–385). Paris: Desclée de Brower.

Ey, H. (ed.) (1966). *L'Inconscient, VIᵉ Colloque de Bonneval*. Paris: Desclée de Brower.

Green, A. (1964). 'Du comportment a la chair: Itineraire de Merleau-Ponty'. *Critique 211*: 1017–1046.

Lacan, J. (1957). 'La psychanalyse et son enseignement'. *Bulletin de la Soci*è*t*é *francaise de Philosophie 2*: 65–101.

Lacan, J. (1978). *The Four Fundamental Concepts of Psychoanalysis*, ed. J.-A. Miller, trans. A. Sheridan. New York: W.W. Norton.

Lacan, J. (1982–3/1961). 'Merleau-Ponty in Memoriam'. *Review of Existential Psychology and Psychiatry 18*(1, 2, 3): 73–82.

Lacan, J. (1988/1954–5). *The Seminar of Jacques Lacan: Book II. The Ego in Freud's Theory and in the Technique of Psychoanalysis*, ed. J.-A. Miller, trans. S. Tomaselli. New York: W. W. Norton.

Lanteri-Laura, G. (1966). 'Les problèmes de l'inconscient et la pensée phénoménologique'. In H. Ey (ed.), *L'Inconscient, VIᵉ Colloque de Bonneval* (pp. 387–408). Paris: Descle de Brower.

Laplanche, J. (1989). *New Foundations for Psychoanalysis*, trans. D. Macey. Oxford: Basil Blackwell.

Laplanche, J. and Pontalis, J.-B. (1973). *The Language of Psychoanalysis*. New York: W. W. Norton.

Loewald, H. (1980). *Papers on Psychoanalysis*. New Haven, CT: Yale University Press.

Merleau-Ponty, M. (1963/1942). *The Structure of Behavior*, trans. A. Fisher. Boston, MA: Beacon Press.

Merleau-Ponty, M. (1962/1945). *Phenomenology of Perception*, trans. C. Smith. London: Routledge & Kegan Paul.

Merleau-Ponty, M. (1964a/1960). 'Man and Adversity'. In *Signs*, trans. R. McCleary (pp. 224–246). Evanston, IL: Northwestern University Press.

Merleau-Ponty, M. (1964b/1960). 'The Philosopoher and his Shadow'. In *Signs*, trans. R. McCleary (pp. 159–182). Evanston, IL: Northwestern University Press.

Merleau-Ponty, M. (1970). *Themes from the lectures at the Collège de France, 1952–1960*, trans. J. O'Neill. Evanston, IL: Northwestern University Press.

Merleau-Ponty, M. (1966). 'Remarks'. In H. Ey (ed.), *L'Inconscient, VIᵉ Colloque de Bonneval* (p. 143). Paris: Desclée de Brower.

Merleau-Ponty, M. (1982–3/1960). 'Phenomenology and psychoanalysis: Preface to Hesnard's *L'oeuvre de* Freud'. *Review of Existential Psychology and Psychiatry 18*(1, 2, 3): 67–72.

Merleau-Ponty, M. (1988/1949–52). Merleau-Ponty á *la Sorbonne: R*ésum*é de cours*. Paris: Cynara.

Merleau-Ponty, M. (1992/1960). 'Merleau-Ponty in person: (An interview with Madeleine Chapal, 1960)'. In H. Silverman and J. Barry (eds), *Texts and Dialogues*, trans. M. Smith (pp. 2–13). Atlantic Heights, NJ: Humanities Press.

Merleau-Ponty, M. (1964b/1961). 'Eye and Mind'. In J. Edie (ed.), *Maurice Merleau-Ponty, The Primacy of Perception* (pp. 159–92). Evanston, IL: Northwestern Univesity Press.

Merleau-Ponty, M. (1968/1964). *The Visible and the Invisible*. C. Lefort (ed.), trans. A. Lingis. Evanston, IL: Northweatern University Press.

Mitchell, S. (1988). *Relational Concepts in Psychoanalysis: An Integration*. Cambridge, MA: Harvard University Press.

Phillips, J. (1988). 'Latency and the Unconscious'. In E. Murphy (ed.), *Phenomenology and the Unconscious: The Sixth Annual Symposium of the Simon Silverman Phenomenology Center* (pp. 31–64). Pittsburgh, PA: The Simon Silverman Phenomenology Center, Duquesne University.

Phillips, J. (1996). 'Lacan and Merleau-Ponty: The Confrontation of Psychoanalysis and Phenomenology.' In D. Pettigrew and F. Raffoul (eds), *Disseminating Lacan* (pp. 69–108). Albany, NY: State University of New York.

Phillips, J. (1999). 'From the Unseen to the Invisible: Merleau-Ponty's Sorbonne Lectures as Preparation for his Later Thought. In D. Olkowski and J. Morley (eds), *Merleau-Ponty, Interiority and Exteriority, Psychic Life and the World* (pp. 69–88). Albany, NY: State University of New York Press.

Pontalis, J.-B. (1971). 'Presence, entre les signes, absence'. *L'Arc 46*: 56–66.

Pontalis, J.-B. (1982–3/1961). 'The problem of the unconscious in Merleau-Ponty's thought'. *Review of Existential Psychology and Psychiatry 18*(1, 2, 3, 4): 67–72.

Ricoeur, P. (1970). *Freud and Philosophy*. New Haven, CT: Yale University Press.

Roudinesco, E. (1990). *Jacques Lacan & Co.*, trans. J. Mehlman. Chicago, IL: Chicago University Press.

Spiegelberg, H. (1982). *The Phenomenological Movement*. The Hague: Martinus Nijhoff.

Notes:

(1.) This and other sections draw on material that appeared in James Phillips (1988).

(2.) Discussed at greater length in James Phillips (1999).

(3.) Discussed at greater length in James Phillips (1996).

(4.) On two occasions after Merleau-Ponty's death Lacan continued his critique. In his tribute to Merleau-Ponty in *Les Temps Modernes*, argued once more that with his concept of the perceptual consciousness, Merleau-Ponty perpetuated a unity of the subject that is not there. 'It is this very presupposition—that there be a place of unity somewhere—that makes us delay our assent' (Lacan 1982–3: 77–8). Lacan summarized his disagreement with Merleau-Ponty with a sharp contrast between two versions of the subject: as constituted primarily through the signifier (Lacan), and as constituted primarily through the body (Merleau-Ponty).

The second occasion was the publication of *The Visible and the Invisible* in 1964. Shortly after the publication, in his seminars of February and March, 1964, Lacan again contrasted Merleau-Ponty's starting point—the perceived world of sensory experience—with his own—human desire and its vicissitudes (1978). At the end of the final discussion of Merleau-Ponty, in the seminar of 11 March 1964, Jacques Miller remarked to Lacan that he had talked of a convergence between Merleau-Ponty and himself regarding the unconscious. Lacan responded: 'I did not say that. I suggested that the few whiffs of the unconscious to be detected in his notes might lead him to pass, let us say, into my field. I'm not at all sure' (1978: 119). Miller finally asked: 'This leads me to ask you if *The Visible and the Invisible* has led you to change anything in the article that you published on Merleau-Ponty in a number of *Les Temps Modernes?*' To which Lacan responded: 'Absolutely nothing' (Lacan 1978: 119).

(5.) See Pietersma (1988) for Merleau-Ponty's relation to Spinoza.

(6.) See Thomas Fuchs (this volume) for a related, but different approach to the unconscious based in large part on Merleau-Ponty. Citing the same quote, he interprets the notion of an unconscious 'not at the bottom of ourselves … but in front of us …' as indicative of the unconscious as bodily memory, as opposed to the monistic view expressed in this chapter of the unconscious as the unseen structure holding the world of flesh together.

(7.) Translation altered. See Merleau-Ponty, M. (1982–3/1960).

James Phillips

Yale Global Mental Health Program, Yales School of Medicine

Oxford Handbooks Online

Abstract and Keywords

Wittgenstein proposes an Aesthetic Model for psychoanalytic interpretation to replace Freud's Causal Model. When we assent to such interpretations of our dreams, for example, he claimed that we are not making inferences based on evidence about the causes of our dreams; instead, our assent is comparable to more dreaming, or 'redreaming' our dreams, as Wittgenstein called it. That assent is part of Freud's criterion of interpretive correctness, which would not be possible if recognizing causal laws were involved, Wittgenstein claims. For causal inferences need to be based on inductive evidence, which, with dreams, is non-existent or unavailable. Since for Freud, interpreting dreams shows how neurotic symptoms, jokes, slips, and much else can be interpreted, i.e. by free association, this non-causal character distinguishes Wittgenstein's Aesthetic Model for psychoanalytic interpretation quite generally—in which dreams and symptoms cannot be assumed to cause the associations we ascribe to them—from Freud's Causal Model.

Keywords: Aesthetic Model, assent, Causal Model, criterion, free association, inductive evidence, interpretive correctness, non-causal, psychoanalytic interpretation, redreaming

Introduction

Wittgenstein wrote nothing on psychoanalysis for publication, but commented on it repeatedly, in written notes made for his own use, as well as in publicly and privately spoken remarks, often of considerable length, reported by others. This heterogeneous body of notes and reports produced over many years has generated considerable interest, since his whole philosophical enterprise from the start bears the signs of Freud's influence, as he freely acknowledged. Like Freud exploring the unconscious, Wittgenstein approached philosophy as a kind of therapy, investigating philosophical problems by uncovering the hidden misconceptions about language enshrined in them, thereby subverting their power to bewitch our minds. This novel approach combined Freud's psychology with a concern for the verification of statements and the dependence of their meaning on the activities in which they are used, which enabled Wittgenstein to examine long-familiar questions in new ways. But in combining that philosophy, according to which "Nothing is hidden" (Wittgenstein 1945–9: §435), with Freud's psychology, for which much of importance in the mind *is* hidden, Wittgenstein's programme was bound to generate tension, which is one of the main subjects of this chapter.

Wittgenstein's attitude to traditional thought could be called reserved, given his greatly reduced dependence on earlier influences or even on the ideas of contemporaries; thus, his response to psychoanalysis was personal and specific in a way that is probably unmatched among philosophers. As a result of that reserve, it is difficult to say where Wittgenstein's take on Freud's work belongs in the conventional dichotomy between hermeneutic and scientific views of psychoanalysis, since unlike most of those who have confronted this dichotomy, Wittgenstein did not stay put on one side of it or the other, sometimes praising Freud's contribution to science but also attacking his explanations of psychological phenomena as unscientific. Such apparent inconsistencies may be merely the result of his uniquely personal 'zig-zag' style of philosophical inquiry, in which wildly different 'thought-experiments' are tried; but the result is that in assessing Wittgenstein's view of psychoanalysis, perfect agreement cannot be expected concerning what these notes and recorded comments and discussions add up to. We cannot even say for sure whether a single, stable view underlies these materials—perhaps one that evolved over time—or whether instead Wittgenstein was content with allowing his observations on psychoanalysis to remain as they now, superficially, appear—isolated insights, surrounded inconclusively by mistakes and changes of mind. In the following account, signs of development in Wittgenstein's views have been emphasized, although this requires ignoring to some degree indications of repeated changes of direction, which are not easily distinguished from that personal style of philosophizing. Obviously, any account based on these notes and reports, and produced under such conditions, must be a tentative and speculative affair.

Although perfect consensus among commentators is absent, on one point something like agreement can be found; it is that whatever his initially favourable opinion of psychoanalysis may have been, Wittgenstein came to be dissatisfied with what he is said to have seen as a "pseudoscience" (Bouveresse 1991: 52; Cioffi 1996: 286) and ended up critical of everything about it. Although the evidence for this very negative view deserves to be reviewed here, a more nuanced conclusion can be supported, according to which, although Wittgenstein often saw Freud as seriously mistaken about the nature of psychoanalysis, his estimation of it did not waver as much as may at first appear.

A Pseudoscience?

The term 'pseudoscience' signifies non-adherence to the norms of scientific research, and is always pejorative, but what more can be said about its meaning is unclear. Speaking of 'pseudoscience' here is troublesome because Wittgenstein seems never to have used the term in reference to psychoanalysis or to anything else, although he certainly expressed disparaging views about the excesses he thought he saw in many of Freud's claims. Nevertheless, we don't generally apply the epithet 'pseudoscience' to a theory merely because some *part* of it (even a large part) deviates from modern scientific standards. After all, it would be absurd to call the Ptolemaic cosmology pseudoscience simply because of its geocentrism and reliance on epicycles, or to dismiss Kepler's astronomy because of its admixture of astrology and numerology. The circular planetary orbits of Copernicus and the luminiferous aether assumed by the Michelson-Morley experiments are further examples; none of these are pseudoscientific at all. What, then, distinguishes genuine science, however riddled with error or imperfectly executed, from its counterfeit?

Attempts to transform 'pseudoscience' from a mere epithet into a determinate term of philosophical critique, and thereby enable objective determination of whether psychoanalysis is scientific or not, begin with Karl Popper, for whom the defining characteristic of a pseudoscience is that its formulations are essentially impossible to falsify and, therefore, immune from empirical testing. Thus, for Popper, given any possible behaviour, e.g. whether someone tries to drown a child or seeks to rescue it from drowning, a psychoanalytic interpretation is available to 'explain' whatever is observed (Popper 1963: 35–7). In this, Popper claims, psychoanalysis differs from the real thing, the claims of which are always open to refutation by observed facts.

Frank Cioffi revised Popper's theory of pseudoscience while also adopting psychoanalysis as a paradigm case of it. Thus, like Popper, Cioffi agrees that psychoanalysis is unfalsifiable; however, for Cioffi, unlike Popper, its unfalsifiability is not what makes it pseudoscientific (Cioffi 1985: 210). For a theory might lead to the discovery of new, empirically testable truths, ones that can be independently verified; but if its proponents

habitually and willfully (even self-deceptively) avoid using methods that aim at such falsification, then the theory would still be pseudoscientific (Cioffi 1970: 116).

The affinities or lack of them between the conceptions of psychoanalysis as a pseudoscience in Popper and Cioffi, on the one hand, and Wittgenstein's criticisms of Freud, on the other, can be left to emerge as we survey Wittgenstein's views; however, two points are constant amidst the shifts in those views. First of all, for Wittgenstein, whatever defects he found in the procedures and assumptions of psychoanalysis, that it might be entirely without cognitive value never seems to have occurred to him. Second, unlike Popper and Cioffi, Wittgenstein never attempted to develop a theory of pseudoscience by which to judge psychoanalysis, so it should not surprise us that he never used the term 'pseudoscience' with regard to psychoanalysis or anything else, as far as we know.

Nevertheless, as Cioffi remarks, Wittgenstein thought psychoanalysis had scientific pretensions that should be abandoned, unlike its scientific critics, who pressed for it to live up to those pretensions (Cioffi 1996: 286). Yet in *The Blue Book*, Wittgenstein also credited Freud with 'newly discovered psychological reactions' (Wittgenstein 1933–4: 57), and referred to Freud's 'extraordinary scientific achievement' in a letter to Norman Malcolm written in 1945 (Malcolm 1958: 39).

That passage in *The Blue Book* does not give examples or specify which psychological reactions Wittgenstein meant. The unwary reader might assume that Wittgenstein was thinking of repression, projection, and transference, for example, but that assumption should give us pause, since all of these, as Freud understood them, involve *unconscious mental causality*, about which Wittgenstein was extremely doubtful when he dictated these words, and about which he remained suspicious long after. Instead, it is possible that the reactions Wittgenstein was thinking of here were ones that appear to be observable and describable without reference to unconscious mental causality—specifically, our reactions when we intuitively grasp the symbolism, especially sexual symbolism, which Freud thought we find in dreams. For dreams can seem to have meaning, even before we think we know what they mean; in this respect, dreams are unlike other phenomena that are interpreted psychoanalytically, which may appear initially to be meaningless. Such symbols in dreams that we do come to accept are often not imposed upon us; they seem instead to arise spontaneously, so we do not need to be persuaded to accept them, as Wittgenstein came to recognize (Wittgenstein 1942–6: 43, 45), although he had earlier thought they were 'impositions' (Wittgenstein 1938: 27). By contrast, Freud's Primal Scene theory, which Wittgenstein called a 'new myth' (Wittgenstein 1942–6: 51) is, Wittgenstein maintained, 'imposed' even when assented to. But whether assented to or not, what, exactly, made the very idea of *unconscious mental causality* so very troublesome for him?

Reasons and Causes

According to Wittgenstein around the time of *The Blue Book*, Freud was wrong to believe that he was the discoverer of the *unconscious causes* of dreams, jokes, slips, and neurotic symptoms. Wittgenstein comments:

> I see a muddle here between a cause and a reason. Being clear why you laugh is not being clear about *a cause*. If it were, then agreement to the analysis given of the joke as explaining why you laugh would not be a means of detecting it. The success of the analysis is supposed to be shown by the person's agreement. There is nothing corresponding to this in physics. Of course, we *can* give *causes* for our laughter, but whether those are in fact the causes is not shown by the person's agreeing that they are. A cause is found experimentally. The psychoanalytic way of finding why a person laughs is analogous to an aesthetic investigation. For the correctness of an aesthetic analysis must be agreement of the person to whom the analysis is given. The difference between a reason and a cause is brought out as follows: the investigation of a reason entails as an essential part one's agreement with it, whereas the investigation of a cause is carried out experimentally … Of course the person who agrees to the reason was not conscious at the time of its being his reason

> (Wittgenstein (1932–5: 39–40).

So it is not the unobservability of unconscious mental activity by itself that leads Wittgenstein to regard any reference to *unconscious mental causality* as suspect; indeed, that supposed unobservability may prove to be less of an impediment than appears. Here, although his criticism still derives from concerns about the verifiability of such causal claims, its focus falls on Freud's claims to have discovered the unconscious mental *causes* of dreams, jokes, slips, and symptoms. According to Wittgenstein, if the criterion of correctness for an interpretation is the assent of the individual subject, then causality cannot be involved. For as he understood matters at this time, legitimate causal claims are the product of repeated experiments (Wittgenstein 1932–5: 39–40), issuing in the formulation of causal laws. By contrast, Freud's supposed unconscious causes are one-time occurrences, and it is only the individual dreamer's or joker's assent, e.g. to the specific interpretation, which is decisive; repeated experiments on other subjects would not be to the point. Thus, for Wittgenstein, what Freud presented as causal explanations were actually reasons.

Wittgenstein would probably find little to criticize in contemporary cognitive science's ascription of unconscious mental causes, which is based overwhelmingly on experimental observations of many subjects, yielding law-like generalizations, where assent of the subjects plays hardly any evidentiary role. Wittgenstein's considerations do not merely subvert the claim of psychoanalysis to uncover unconscious mental causes; his view also undermines the *necessity* of ever invoking 'the Unconscious' at all, even if unconscious

ideas and wishes are allowed (Bouveresse 1991: ch. 2). Thus, when we laugh at a joke, for example, psychoanalysis does not give us the unconscious cause of our laughing, as Freud supposed; instead:

> … it is a way of speaking to say the reason was subconscious. It may be expedient to speak in this way, but the subconscious is a hypothetical entity which gets its meaning from the verification these propositions have. What Freud says about the subconscious sounds like science, but in fact it is just a *means of representation*. New regions of the soul have not been discovered, as his writings suggest

> (Wittgenstein 1932–5: 40).

The 'means of representation' Wittgenstein speaks of here is capable of becoming what he elsewhere calls a 'mythology' when it is erroneously viewed as something represented, or as "persuasion" when others, under pressure, accept the conflation. For Wittgenstein, quite generally, when you assent to a psychological explanation involving unconscious reasons, '… you could have been persuaded of something different', if other means had been employed (Wittgenstein 1938: 27). However, there is no need to assume the existence of 'the Unconscious'—doing so is merely one 'means of representation' among many, where none is the single, uniquely correct way to represent the observable facts. Wittgenstein does not pause here to consider how he knows that Freud had *not* discovered 'new regions of the soul'—as if he knew what conditions for such a discovery Freud's evidence had failed to satisfy, and what evidence would have satisfied them. How does he know when something is merely a 'means of representation' and nothing more? Perhaps a means of representation might *become* something else—as atoms and molecules, which, until Einstein, were thought to be merely hypothetical entities, later came to be understood to be real. And if unconscious reasons are unproblematic, why is 'the Unconscious' different? For if poor people or baroque buildings exist, references to 'the poor' or to 'the baroque' should pose no challenge. These questions are not as peripheral as might appear, since they bear upon our earlier tentative rejection of repression, projection, and transference, as instances of what Wittgenstein could have meant when he spoke of Freud's 'newly discovered psychological reactions'. That earlier rejection seems to be undermined by the following passage from *The Blue Book*, dictated at about the same time as the lectures quoted above, which shows him accommodating 'unconscious thoughts, acts of volition, etc.', although in a very mitigated way:

> It might be found practical to call a certain state of decay in a tooth, not accompanied by what we commonly call toothache, 'unconscious toothache' and to use in such a case the expression we have a toothache, but don't know it. It is in just this sense that psychoanalysis talks of unconscious thoughts, acts of volition, etc. Now is it wrong in this sense to say that I have a toothache but don't know it? There is nothing wrong about it, as it is just a new terminology and can at any time be retranslated into ordinary language

> (Wittgenstein 1933–4: 22–3).

The suggestion here is that for Wittgenstein, against our earlier hesitation, there could be observable reactions that would allow ascribing unconscious thoughts and acts of volition, even if there would be no *necessity* in doing so. Repression, for example, might be conceived, roughly, as trying to forget an untoward thought or wish, by habitually switching attention to something else whenever it comes to mind, while preventing other thoughts or wishes from bringing it to mind by habitually shunting them aside whenever they verge on doing so. Being habitual, the 'switching' and 'shunting' of attention might escape notice by the subject, resulting in a motivated inability to attend to the untoward thought. Even if the specific unconscious thought or wish being repressed by an individual cannot be identified by observers, that some idea or wish is being repressed might be, based on such evidence. Thus, at least some psychoanalytic references to repression, so long as they can be 'retranslated into ordinary language' in this way, might be unobjectionable, for Wittgenstein.

It is strange for Wittgenstein to claim that 'unconscious toothache' illustrates something about psychoanalysis, since unlike in psychoanalysis, the assent of the tooth's bearer is not the criterion for ascribing unconscious toothache; if there is tooth decay in a subject's tooth but no felt pain, there is what may be called 'unconscious toothache', period. Indeed, if and when the tooth's bearer knows of the decay and feels no pain, they would then be fully conscious of experiencing unconscious toothache. Whether such a thing could be true of the sort of unconscious mental states Freud discussed, that is, whether the subject of a particular unconscious thought or act of volition could at the same time be conscious of being such a subject remains to be seen. Perhaps we are tripping here over the limitations imposed by relying exclusively on questions of verifiability, very narrowly conceived, with which to investigate the meaning of psychoanalytic claims. For his comments on those claims seem to gravitate toward saying either that psychoanalytic interpretations are unverifiable, period (inasmuch as they are not based on repeatable experiments, and so cannot ascribe causes), or that the sole verification of their correctness is the assent of the individual subject.

Aesthetics as the Model

An alternative conceptualization of Freud's observations, one closer to the truth, according to Wittgenstein, was available in Aesthetics, which he regarded as a branch of Ethics (Wittgenstein 1929: 38), wherein observable similarities and differences are to be found, without experiments, inductive inferences, or inferences of any kind (Wittgenstein 1938: 18), and without having to refer to unobservable activities in an Unconscious Mind. Absent the need to prove causality, the assent criterion for interpretations could now be taken as sufficient. Then a dream image, say, of a flowering tree branch could be a sexual symbol—without concern about what, if anything, really caused the dream in the first place. Such an understanding of dream symbolism would solve the puzzle concerning what Wittgenstein was thinking of when he praised Freud for his newly discovered

psychological reactions while at the same time emphasizing the absence of any need to resort to unconscious mental causality to describe or explain dream reports and our intuitive understanding of their meanings. As Wittgenstein remarked to Rush Rhees in 1942:

> Freud in his analysis provides explanations which many people are inclined to accept. He emphasizes that people are *dis*-inclined to accept them. But if the explanation is one which people are disinclined to accept, it is highly probable that it is also one which they are *inclined* to accept. And this is what Freud had actually brought out. Take Freud's view that anxiety is always a repetition in some way of the anxiety we felt at birth. He does not establish this by reference to evidence—for he could not do so. But it is an idea which has a marked attraction. It has the attraction which mythological explanations have, explanations which say that this is all a repetition of something that has happened before
>
> (Wittgenstein 1942–6: 43).

This is evidence that what Wittgenstein meant when he referred to Freud's 'newly discovered psychological reactions' were the observable *inclinations* to accept accounts of a sort that are not susceptible to ordinary or scientific inquiry, i.e. ideas which have a marked attraction but whose attraction resist explanation. Such inclinations and attractions might be 'what Freud had actually brought out'. These observable reactions, without unconscious mental activity as causes, favour the Aesthetic approach.

But Wittgenstein was forced to question his Aesthetic assumptions when Bertrand Russell's essay, 'The Limits of Empiricism', with its new ideas about causality appeared (Russell 1935–6); for his notes on causality (Wittgenstein 1937), were apparently written with Russell's essay in mind.

Causality

In that essay, Russell argued that it must be possible to *perceive* causal relations, that is, e.g. to *see, hear, or feel* one thing *producing* another. Such awareness of one thing causing another is necessary before we can discover causality through repeated experience, according to Russell. As an example he offered the fact that:

> When I am hurt and cry out, I can perceive not only the hurt and the cry, but the fact that the one 'produces' the other
>
> (Russell 1935–6: 149).

This claim is a direct rejection of Wittgenstein's assumption that knowledge of causality always arises inductively, that is, by observing regularities in repeated experiments leading to the formulation of causal laws. Instead, for Russell, we can perceive causes at

work in individual cases; so Russell says: '"Cause" accordingly, must mean something other than "invariable antecedent …"' (Russell 1935–6: 137). For Russell, as Wittgenstein paraphrases him:

> … we don't always arrive at the proposition 'this is the cause of that' by observing regular sequences. One has first to recognize certain causes intuitively—I am immediately aware that this is the cause

> (Wittgenstein 1937: 408).

So our ability to engage in *inductive* inferences rests on our *prior* ability to 'intuitively' perceive some causes 'immediately'. However, the notion of *immediate awareness* here is suspicious, for Wittgenstein, since it might be taken to imply that there is a kind of *infallibility* attached to that immediate awareness. Wittgenstein asks:

> Don't we recognize immediately that the pain is produced by the blow we have received? Isn't this the cause, and can there be any doubt about it?—But isn't it quite possible to suppose that in certain cases we are deceived about this? And later recognize the deception?

> (Wittgenstein 1937: 373).

To which he replies:

> Certainly there is in such cases a genuine experience which can be called 'experience of the cause'. But not because it infallibly shows us the cause; rather because one root of the cause-effect language-game is to be found here, in our looking out for a cause

> (Wittgenstein 1937: 373).[1]

After reading Russell on causality, Wittgenstein could no longer simply deny the coherence of someone claiming to be aware of their ideas and wishes causing their individual dreams. Nor could a dreamer experiencing their free associations to a dream as what focusing attention on the dream 'made them think of' be so easily dismissed. The dichotomy between reasons and causes, treated by many as a truism before Russell, became a subject for philosophical debate thereafter, quite apart from its relevance to psychoanalytic interpretation, and it remains so today. Elizabeth Anscombe's distinction between reasons and causes (Anscombe 1957: §§9–11, §14), is less absolute than Wittgenstein's dichotomy; building on her work, and aided by her separation of causality from its deterministic distortions (Anscombe 1971), Donald Davidson's view of reasons as essentially causal gained widespread acceptance (Davidson 1963).

It might at first appear that once Wittgenstein had recognized the reality of perceptions of causality, the dichotomy between his own Aesthetic Model and Freud's scientific approach would have to be reconsidered, since his model had assumed the impossibility of such non-inductive, individual perceptions, whereas Freud had raised no problem

about them. Yet despite his new recognition of their existence, Wittgenstein seemed not to realize their subversive implications for the dichotomy. After all, acknowledging the reality of individual perceptions of causality and recognizing free association as an instance of it, may seem too disparate to immediately connect. The 'freedom' we experience when we free associate to a dream means that then the mind has been temporarily freed of distractions from the dream as fully as possible, and this requires not only the voluntary suspension of action, and of the critical faculty needed for action, but also of thinking about anything to do with action. Then, when one focuses attention on a dream or any of its elements, the thoughts and feelings which follow upon this 'emptying' of the mind are commonly experienced as having been evoked by the dream. Those potentially new thoughts and feelings answer the question, 'What does this dream (or dream element) make you think of?', which cannot be assumed if the question is asked when the freedom that free association provides is absent.

Perhaps he saw that even if free associations to a dream, experienced as the dream's causal efficacy when the dreamer's attention is focused on it, were not really exceptions to our basic beliefs about causality, it did not follow that they were completely unobjectionable. All that Russell had shown is that such causal experiences are not, prima facie, incoherent, as Wittgenstein had believed. After all, he did not yet know whether free associations were similar enough to uncontroversial experiences of causality for them to be entirely problem-free, nor did he know what kinds of evidence would support unconscious ideas and wishes as causes, for example, of dreams or symptoms. For whatever reasons, Wittgenstein's initial response to Russell was not an attempt at integrating Freud's Causal Model with his own Aesthetic, i.e. non-causal, non-inferential approach. On the contrary, in his lectures on Aesthetics, in the summer of 1938 (Wittgenstein 1938: 1–36), Wittgenstein's starting point was the assumed contrast between the two, although glimmers of interest in reconciling them appear.

When he returned to discussing psychoanalysis (Wittgenstein 1942–6), in conversations alone with Rush Rhees (among the most knowledgeable of his students about Freud's writings, to judge from the notes to the lectures on Aesthetics), Wittgenstein at last abandoned the stance he had maintained up to then, of using psychoanalysis merely to illustrate one or another philosophical point, as in the exploration of the assent criterion in those 1938 lectures. Now, for the first time, he approaches psychoanalysis as a new and independent philosophical subject in its own right. As such, these conversations are among the earliest investigations—if they are not the very first—by anyone that have come down to us.

Dreams and Their Associations

The first conversation focuses on free association and how it can be thought to lead to 'the right analysis' of a dream, given that 'Freud never shows how we know where to stop', and that 'it does not seem to be a matter of evidence' (Wittgenstein 1942–6: 42). This implies that interpretations are not inferences, so gleaning clues from many associations is not involved, in Wittgenstein's view. Wittgenstein does not question his own assumption that the right analysis of a dream is to be found in some one thought that occurs, without inference, in the course of the dreamer's free associating, and so that is where the associating should stop, if we only knew where that was. However, the assumption is troubling for two reasons; it seems to tilt, in a question-begging way, in favour of the Aesthetic Model and against the Causal Model since, even if not inferred, causality is arguably multiconditional (Anscombe 1971), and it seems arbitrary to assume that all of the conditions producing the dream will be found in any one association to it. Admittedly, among the associations to a dream Freud is prepared to single out a thought 'which is represented in [the dream's] content and is indispensible for its interpretation' (Freud 1900: 281); but this does not imply that such a thought by itself provides the interpretation. Besides, Wittgenstein's apparent conflating of such an association with the dream's correct interpretation, implies that for Wittgenstein—and contrary to Freud—the process of free association, all by itself, generates interpretations the way dreams generate associations, that is, without the use of the dreamer's 'critical faculty' (Freud 1900: 102–3), which is to say, non-inferentially.

Wittgenstein acknowledges that for Freud, independent evidence supporting a dream interpretation could be found in 'memories brought to light in analysis', i.e. memories not previously recalled (Freud 1937: 262); these might support the interpretations that appear to have prompted them. But that idea is rejected here, although the reason given, i.e. concern about the analyst's role in producing those recollections, seems less than conclusive (Wittgenstein 1942–6: 43). Here, Wittgenstein sweeps aside the possibility that some of those recollections are verifiably true; however, this idea reappears in the next year's conversations (Wittgenstein 1942–6: 46) with a different response. The remainder of the conversation in the summer of 1942 is devoted to the idea of a dream language, for which there is evidence, as Wittgenstein initially agrees, although doubts about it immediately arise:

> Consider the difficulty that if a symbol in a dream is not understood, it does not seem to be a symbol at all. So why call it one? But suppose I have a dream and accept a certain interpretation of it. *Then*—when I superimpose the interpretation on the dream—I can say 'Oh yes, the table obviously corresponds to the woman, this to that etc'

> (Wittgenstein 1942–6: 44).

Here, the dream initially seems to be devoid of symbolic significance, and its contents appear to become symbolic only after the interpretation is accepted and 'superimposed' onto it. However, by the following year, 1943, Wittgenstein seems to reconsider his ideas about psychoanalytic interpretations as 'impositions':

> It is characteristic of dreams that often they seem to the dreamer to call for an interpretation. One is hardly ever inclined to write down a day dream, or recount it to someone else, or to ask 'What does it mean?' But dreams do seem to have something puzzling and in a special way interesting about them—so that we want an interpretation of them. (They were often regarded as messages)
>
> (Wittgenstein 1942–6: 45).

Here, Wittgenstein recognizes that the dreamer wants an interpretation of the dream before one is offered. But he does not completely abandon all aspects of the 'imposition' model; for once a dream is interpreted: '… we might say that it is fitted into a context in which it ceases to be puzzling. In a sense the dreamer re-dreams his dream in surroundings such that its aspect changes' (Wittgenstein 1942–6: 45), a formulation that he does not abandon even when the interpretation prompts recollections, sometimes truthful ones, from early childhood: '(All this is connected with what was said about dreaming the dream over again. It still belongs to the dream, in a way)' (Wittgenstein 1942–6: 46).

Those new recollections do not *necessarily* count as *evidence* for the correctness of the interpretation that prompted them; but neither is there any impediment to adopting an explicitly *causal* framework for incorporating the occurrence of such recollections into an explanatory account of the dream; for immediately following the last quoted remarks, he goes on:

> On the other hand, one might form an hypothesis. On reading the report of the dream, one might predict that the dreamer can be brought to recall such and such memories. And this hypothesis might or might not be verified. This might be called a scientific treatment of the dream
>
> (Wittgenstein 1942–6: 46).

Here, not for the first time (compare Wittgenstein 1938: 18, n3), Wittgenstein recognizes the faint possibility of integrating the Aesthetic and Scientific approaches to psychoanalytic interpretation, instead of contrasting them dichotomously, as he usually does.

But why treat a dream interpretation as 'dreaming the dream over again' merely because, when interpreted, the dream's aspect changes? Wittgenstein's reason seems to be that not only has the dream's aspect changed when it is interpreted, but it has done so with no

provable cognitive gain—as a result of which the interpreted dream has the cognitive status of a new dream, and nothing more.

Yet there is nothing unusual about aspect change in interpreting dreams, compared to other kinds of interpretation; quite generally, in any inquiry, say, a criminal investigation, individual pieces of evidence change aspect as they are connected to different theories of the crime. But even if we treat each new interpretation of a dream, literally, as *redreaming the dream*, there is no reason why a second dream cannot be evidence affecting the interpretation of the original dream. Besides, given repeated memories of a dream (perhaps separately recorded), a dream might be expected to change over time, regardless of whether it has been interpreted or reflected upon. That is, the dream's changes in aspect might result, not only from its having been interpreted or reflected upon, but also from an instability in dreams quite generally, a tendency to drift, like other remembered mental states on which we do not act (Levy 2004: 183–5).

Thus far, Wittgenstein has merely questioned reliance on free association in interpreting dreams given our limited knowledge of the connections between free associations and dreams, since free association 'is likely to be conditioned by everything that is going on about you and within you' (Wittgenstein 1942–6: 47)—which of course could also be said about the dream itself. However, the categorical denial of free association as yielding evidence of the causes of dreams is not yet the issue; thus far, it is only the *degree of reliability* of our perceptions of causality linking dreams and free associations that is in question, i.e. the low probability of that linkage really being as we perceive it, and the small possibility of correcting our misperceptions of it. But Wittgenstein's implicit point here seems to be the absence of *criteria* for distinguishing associations that are genuinely caused by focusing attention on dreams from those that are not—for all such associations are, or can be, experienced as dream-induced, but it is doubtful that all really are.

Free association receives no direct attention for the remainder of the conversations in 1943 (Wittgenstein 1942–6: 47–50), but when Rhees's notes resume, in 1946, it is again discussed. Now, more radical doubts enter concerning whether free associations to a dream can be presumed to provide reliable evidence concerning the dream's cause. The dreamer's free associations to the dream were supposed to imply unconscious ideas and wishes from which the dream's cause could be inferred. But that inference breaks down if something other than the dreamer's attending to the dream actually caused the free associations, unbeknownst to the dreamer; for then, even if the dreamer's free associations do support ascribing certain unconscious ideas and wishes to the dreamer, when those free associations turn out not to have been caused by the dreamer's attention to the dream, the inference collapses, and nothing about the cause of the dream can be inferred from the free associations the dreamer produced. Even when experienced as perceptions of causality, free associations are still fallible, which subverts the project of psychoanalytic interpretation based on free association, for Wittgenstein.

In each of the three arguments that follow, the apparent causal links between focusing on dreams and the dreamer's subsequent free associations are illusory, even though free associations are (fallible) *experiences of* such causal links. As a result, a dream's cause, supposed to be inferable from the unconscious ideas and wishes implied by the associations the attention to the dream apparently produced, proves not to be inferable from the dream after all. Since, by hypothesis, in all three arguments, the dream activity plays no part in producing those associations, the dreamer's associations provides no real evidence concerning the cause of the dream. Three arguments for this are presented:

[I] You could take *my* dream report, and by free association, 'arrive at the same results' (Wittgenstein 1942–6: 50) as those I reach—although it was not your dream. Here, my dream *activity*, as Wittgenstein calls it, plays no causal role in generating *your* free associations; it was the *report* of my dream that did. The argument here is supposed to show the lack of connection between my dream activity and your associative process, which proceeds as if my dream activity caused the associations you produced, when it did not. The point is that with or without my actual dream activity, the separate associations and interpretations produced by us are alike; so my dream activity is not the cause of either.

The argument resembles a *defender* of free association as revealer of causes of dreams might make to justify belief in a shared dream symbolism. Wittgenstein wavered back and forth about dream symbolism, but ultimately rejected it, partly because such a symbolism could not rise to the level of a dream *language* (Wittgenstein 1942–6: 48), i.e. one that would enable translations into and out of waking language, and partly because even if such a symbolism could account for *some* dream elements, it cannot be assumed to be capable of explaining everything in a dream (Wittgenstein 1942–6: 49).

Objection: 1. Your associations to my dream report, which are all experiences of causality, are related, *indirectly,* by way of that report, to my dream activity; but even my own associations are not *directly* related, causally, to my dream *activity*. My associations are related to that activity, i.e. to the dream itself, only indirectly, *by way of my memory* of it. So both sets of associations, yours and mine, are indirectly related to the dream activity itself, although in different ways. 2. That a symbolism is not a language is not necessarily a defect, nor is incompleteness in an explanation a defect either, unless it purports to be complete, which a dream interpretation does not do.

[II] A thing you are preoccupied with, sex, for instance, will lead your associations 'finally and inevitably' back to itself (Wittgenstein 1942–6: 51), no matter where you began. Your starting point, e.g. focusing attention on a dream, might play no actual role in causing your associations to go where they do; possibly, they were caused by the preoccupation (here, sex), and not by your attention to your dream. Earlier (Wittgenstein 1942–6: 46–7), Wittgenstein had argued that a 'whole host of circumstances' could hijack your free associations. Here, in [II], just one such constant distraction, e.g. sex, will lead 'back to that same theme' (Wittgenstein 1942–6: 51), and away from wherever your associations

might otherwise have gone. So there is no warrant for ascribing the cause of the associations to the attention paid to the dream, Wittgenstein claims, since a credible alternative cause, i.e. the preoccupation, exists.

Objection: If the dream is assumed to be subject to the same kinds of causes as is everything else you experience, then the dream as well as the dreamer's associations would be vulnerable to influence by the preoccupation, in which case the preoccupation would not be a distraction from the dream. Unless the dream is assumed to be impervious to causal influences—a question-begging assumption—then either the preoccupation influenced both the dream and the associations separately, or it influenced the associations indirectly, *by way of* the dream. For the preoccupation to be a distraction, *either* the dream *or* the preoccupation are causally linked to the associations *but not both* —a false dichotomy.

[III] After associating to one of your dreams and interpreting it, if you then turn to the objects you find on a table and associate to them, the same pattern might appear in both sets of associations, which seems—absurdly—to speak in favour of interpreting the objects in the same manner as the dream, even though your dream activity did not place those objects on the table. So the objects did not really cause your pattern-bearing associations; but why then assume that your dream activity produced the pattern-bearing associations when you free-associated to the dream? The mere presence of the pattern-bearing associations alone tells us nothing about the cause of the objects' presence or of the dream.

In this account of the argument, [III] is a continuation of [II], since it depends upon the dreamer's unnoticed *persistence of mind* when being led back to the same theme on which the dreamer had been dwelling shortly before, i.e. to the hypothesized dream interpretation. That is why Wittgenstein says of the objects on the table that 'you could find that they all could be connected in a pattern like that' (Wittgenstein 1942–6: 51), i.e. *like the pattern in the associations supporting the dream interpretation*. On this reading, all three arguments challenge the assumption that attending to a dream actually causes one's associations to it; and then those associations do not lead to *the dream's cause*, since the unconscious ideas and wishes those associations imply do not reveal the dream's cause if those associations were not caused by focusing attention on the dream.

Objection: Wittgenstein seeks to reveal an absurdity in employing free association to determine a dream's cause. But the argument rests on the supposed psychoanalytic assumption that the cause is revealed in a single free association (compare Wittgenstein 1942–6: 42), whose pattern here reappears when you associate to objects on a table. Supposedly, psychoanalysis involves interpreting the same pattern in the same way, regardless of where it appears, as the sexual symbolism in dreams seems to warrant, which would lead to the absurdity of interpreting your associations to the objects on the table as if you had dreamt them. But requiring that a single pattern-bearing association identifies the correct interpretation is questionable, and ignores the role of *convergence*

among many associations in arriving at an interpretation. With convergence in play, and without that questionable requirement, Wittgenstein's (desired) absurd implication fails.[2]

On this reading, Wittgenstein's purpose in all three sub-arguments is merely to show that free association yields no good evidence from which the causes of dreams can be inferred; that is, estimating the cognitive value of free association in all its aspects is not under discussion. So just because free association to a dream 'does not explain why the dream occurred', this does not imply that free association has no cognitive value at all, since one 'may be able to discover certain things about oneself' by its means (Wittgenstein 1942–6: 51), and such discoveries might be momentous. Thus, in a private notebook entry dated 1939, he writes:

> In a way having oneself psychoanalyzed is like eating from the tree of knowledge. The knowledge acquired sets us (new) ethical problems; but contributes nothing to their solution
>
> (Wittgenstein 1977: 34).

which shows that he thought psychoanalysis, and therefore, free association, enables genuine discoveries about oneself.

Wittgenstein's arguments in [I], [II], and [III] then, challenge the psychoanalytic view that free associations provide evidence of dreamers' unconscious ideas and wishes, and that from those unconscious ideas and wishes their dreams' causes can be inferred. For him, relying on free association in this way assumes either that the perceptions of causality comprising associations to dreams are *infallible*, or else that we have criteria—or can discover them—for distinguishing veridical perceptions of causality among those associations from the rest. For without one or the other of these assumptions, the reliance on free association for evidence of causality would be unjustified, in his view. His arguments are meant to show that neither assumption is justified, for no perceptions of causality are infallible, and criteria for perceptions of causality among these associations are not forthcoming.

Blocking our search for those criteria are two seemingly unavoidable conditions; one is the *instability* of dreams, i.e. their aspect changes and their lack of any firm identity—and the fact that we only know about our own dreams by way of our unverifiable memories of the 'original' dream activity. This instability sets dreams apart from causes in ordinary perceptions of causality. The second condition is the *opacity* of associations to dreams, i.e. we cannot distinguish associations whose sources are truly the dreamer's focusing on their dreams, from those that originate elsewhere; for both kinds of associations are experienced alike as effects of dreamers focusing attention on their dreams. The upshot is that however much it may seem to us that focusing on our dreams causes our associations to them, there seems to be no empirical means of determining when, if ever, this really is so.

So even granting Russell's claim concerning the existence of individual causal perceptions, Wittgenstein still has reasons to resist treating free associations *to dreams* as similar enough to other kinds of causal perceptions, i.e. ones whose veracity is unquestioned. But would comparable resistance be warranted against treating individuals' associations to their neurotic symptoms, jokes, or slips as veridical causal perceptions? In those cases, isolation from action and total reliance on memory, which seem to be implicated in the instability of dreams and the opacity of their associations, are not obvious impediments. Still, Wittgenstein seems to be holding evidence of causality in dream associations to a stricter standard, beyond anything assumed in other, unquestioned perceptions of causality; for him, searching psychoanalytically for the causes of associations to dreams imposes upon the analyst the burden of proof in a practically impossible task, i.e. proving of each and every possible cause of an association aside from the dream, that none had been operative in the association's production, a demand that could not be justified merely by referring to isolation from action and total reliance on memory alone (Wittgenstein 1942–6: 46–7). But if this unreasonable burden of proof is not to be shouldered, how else can psychoanalysis claim plausibility, let alone credibility, in treating some associations to dreams as veridical causal perceptions?

Convergence, Ambiguity, and 'Seeing-As'

Up to this point, Wittgenstein has assumed that for psychoanalysis, the *correct* explanation of a dream, i.e. its cause, is to be found, non-inferentially, in some one 'tell-all' association (Wittgenstein 1942–6: 42, 47, 51). That is, unlike Freud, he had ignored the possibility that even if no one association to a dream by itself reveals the cause of the dream's content, the dreamer's associations, taken together, might *converge*, thereby providing evidence of the dreamer's unconscious ideas and wishes, from which the cause of the dream's content could be inferred; and that might remain true even if each separate association is a fallible, even uncertain, perception of causality. For fallibility and uncertainty differ in that the latter admits of degrees, unlike the former; whereas many gradations in certainty and uncertainty exist, a perception or belief is either fallible or infallible, with nothing between; and even very certain beliefs remain fallible. Thus, a set of fallible perceptions, by themselves, no matter how numerous, do not generally warrant an infallible judgement (but might warrant a very certain one), whereas a large number of individually less-than-certain converging perceptions, especially ones from independent sources, could, together, justify a very certain judgement.

Introducing convergence among associations as evidence, whether for the meaning of a dream or for its cause, suggests that even if the dreamer's assent to an interpretation is taken to be a *criterion* of its correctness, as the Aesthetic Model assumes, that assent need not be regarded as indefeasible. In practice, that apparent infallibility can be replaced by the (not unconditional) deference that dreamers' greater familiarity with

their own lives normally warrants, and then causal explanations of dreams would more easily be seen as based on converging evidence from multiple sources, including converging associations, in which case none needs to be infallible (Levy 1996: 19–23).

Recognizing convergence among associations as evidence also suggests that in judging the reliability of free association as a method of inquiry, the general fallibility of dreamers' beliefs about the causes of individual associations could be less important than Wittgenstein assumed. A dreamer's assent to an offered interpretation, or inclination to assent to one, or support for one that they themselves produced, might be evidence relevant to an interpretation, even if the associations on which it rests are of doubtful origin—indeed, regardless of whether those perceptions of causality are veridical. For the relevance of those associations as evidence in interpreting the dream might depend simply on their being *seen as* veridical perceptions of causality by the dreamer. So interpreting a dream, by focusing exclusively on the veridicality (or lack of it) of the perceptions of causality comprising our free associations to it, while ignoring the dreamer's belief that those perceptions of causality are veridical, could mean throwing out potentially valuable evidence.

However, convergence in practice brings with it the possibility of *ambiguity* in the outcome, since a set of associations that converge to a point could converge to one or more other points as well—which cannot occur in psychoanalytic dream interpretation on the Aesthetic Model's assumption about what Freud claims (Wittgenstein 1942–6: 42, 47, 51). For Wittgenstein, psychoanalysis assumes that the correct interpretation of a dream is to be found in one association to it—an assumption that would be hard to justify unless Wittgenstein also assumed that Freud partly shared his own earlier picture theory of meaning as expounded in the *Tractatus* (Wittgenstein 1921). According to that theory, abandoned by him by the time of these discussions of psychoanalysis, if a dream is to be meaningful at all, it must share the same 'logical multiplicity' (Wittgenstein 1921: §4.04) with its meaning, i.e. with the state of affairs it depicts—and with the single association which, according to Wittgenstein, psychoanalysis assumes to be its correct interpretation. For without that background assumption, the expectation that one association must always simply *be* the dream's meaning would be very hard to explain. So the Aesthetic Model rules out consideration of convergence, a priori, and ignores the sort of ambiguity needing interpretation that convergent associations can produce (Wittgenstein 1921: 3).

Now such ambiguous perceptions, produced by convergence, are instances of *seeing-as* phenomena, a much-discussed topic in Wittgenstein's later writing (Wittgenstein 1945–9; 1946–9). For instance, we *perceive* our associations to a dream *as* caused by it, although we can also perceive them as not so caused—comparable to the way the ambiguous duck–rabbit drawing can be seen in two different ways. However, in many of the other cases of ambiguity that Wittgenstein considers, the ambiguity is eliminated (or can be) even if we often do not know how we do it; by contrast, in perceiving a dream as the cause—or not—of our associations to it, we seem to be unable to eliminate the ambiguity with any

finality, just as we cannot eliminate the ambiguity in a duck–rabbit drawing if its creator intended to draw something ambiguous.

Indeed, some ambiguous associations continue to be *perceived* by their dreamers *as* induced by focusing attentively on the dream, yet perhaps because of its unwelcome implications, those dreamers doubt the causal link while still feeling its force (Wittgenstein 1942–6: 50–1). This immunity or resistance to final refutation marks off *our inclination* to believe that certain associations were caused by focusing attention on the dream. For in some cases, and at certain times, even when particular associations are not believed by the dreamer to be caused by their attending to the dream, those associations continue to seem to them to be caused by the dream, comparable to the way a spoon in a clear glass of water 'looks bent', even though the viewer does not believe it really is.

Such dream associations, which remain causally ambiguous to the dreamer, seem to provide a new kind of evidence beyond the shared, relatively unambiguous dream-symbols recognized by the Aesthetics Model. Like them, these ambiguous causal perceptions also result in spontaneous verbal reactions to dreams that are apparently inexplicable by rational or common sense considerations; the dreamer does not understand why those associations continue to seem to be caused by attending to the dream, even while doubting or denying that they are. (Such a dreamer resembles the one in [III], who, after associating to objects on a table, is tempted to interpret them as if they were a dream).

Those dream associations can plausibly be seen as evidence of an *individual* (i.e. not necessarily shared) dream symbolism, but one without solipsistic (Private Language) implications, since the test for immunity from correction in the bent-spoon kind of case is not private. So the version of free association that results from weighing the dreamer's perceptions of causality, including ambiguous ones, in investigating a dream's causes resembles the Aesthetic Model's version of free association. In both models, uncovering what the dream means to the individual dreamer is the point of the endeavour, where the individual dream symbolism provides objective evidence of such meaning, apart from the dreamer's acceptance.

Conclusion

Wittgenstein's view of psychoanalysis is very far from being the total condemnation that Cioffi supposed, and as it still appears to many readers. Unlike Wittgenstein, Cioffi found little or nothing in psychoanalysis worth salvaging—hence it is no surprise that he viewed it simply as a pseudoscience. Yet when we, today, get past the dichotomies in which much of Wittgenstein's thinking about psychoanalysis was embedded, we can see why Rush Rhees was convinced that their conversations on it show Wittgenstein 'trying to separate what is valuable in Freud from that "way of thinking" which he wanted to

combat' (Wittgenstein 1942–6: 41). But whether he ever fully achieved that separation, and what it might have consisted in, we do not know.[3]

References

Anscombe, G. E. M. (1957). *Intention*. Oxford: Blackwell.

Anscombe, G. E. M. (1971). 'Causality and Determination'. In *Collected Papers*, Vol. 2, *Metaphysics and the Philosophy of Mind* (pp. 133–147). Oxford: Blackwell.

Bouveresse, J. (1991). *Wittgenstein Reads Freud: The Myth of the Unconscious*, trans. C. Cosman. Princeton, NJ: Princeton University Press.

Cioffi, F. (1969). 'Wittgenstein's Freud'. In *Freud and the Question of Pseudoscience*. Chicago, IL: Open Court.

Cioffi, F. (1970). 'Freud and the Idea of a Pseudoscience'. In *Freud and the Question of Pseudoscience*. Chicago, IL: Open Court.

Cioffi, F. (1985). 'Psychoanalysis, Pseudoscience, and Testability'. In *Freud and the Question of Pseudoscience*. Chicago, IL: Open Court.

Cioffi, F. (1996). 'Through the Psychoanalytoscope'. In *Freud and the Question of Pseudoscience*. Chicago, IL: Open Court.

Davidson, D. (1963). 'Actions, Reasons and Causes'. In *Essays on Actions and Events*. Oxford: Oxford University Press.

Drury, M. O'C. (1984). 'Conversations with Wittgenstein' (ed. R. Rhees). In F. A. Flowers III and Ian Ground (eds), *Portraits of Wittgenstein*, Vol. 2 (pp. 828–829). London: Bloomsbury Academic.

Freud, S. (1900). *The Interpretation of Dreams*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 4. London: Hogarth Press.

Freud, S. (1937). 'Constructions in Analysis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 23. London: Hogarth Press.

Harcourt, E. (2017). 'Wittgenstein and Psychoanalysis'. In Hans-Johann Glock and John Hyman (eds), *A Companion to Wittgenstein*. Oxford: Wiley-Blackwell.

Holmes, J. (2010). 'Charles Rycroft's Contribution to Contemporary Psychoanalytic Psychotherapy', *The American Journal of Psychoanalysis 70*: 180–192.

Levy, D. (1996). 'Wittgenstein's Critique of Psychoanalysis'. In *Freud Among the Philosophers: the Psychoanalytic Unconscious and Its Philosophical Critics*. New Haven, CT: Yale University Press.

Levy, D. (2004). 'Incompleteness and Experimental Untestability in Psychoanalysis'. In Jon Mills (ed.), *Psychoanalysis At The Limit; Epistemology, Mind, and the Question of Science*. Albany, NY: State University of New York Press.

Malcolm, N. (1958). *Ludwig Wittgenstein: A Memoir*, 2nd edn 2001. Oxford: Oxford University Press.

Moore, G. E. (1954). 'Wittgenstein's Lectures in 1930–33'. In *Ludwig Wittgenstein: Philosophical Occasions 1912–1951*, eds J. Klagge and A. Norman (pp. 45–114). Indianapolis, IN/Cambridge: Hackett.

Popper, K. (1963). *Conjectures and Refutations*. Oxford: Routledge & Kegan Paul.

Prum, R. O. (2017). *The Evolution of Beauty*. New York: Doubleday.

Russell, B. (1935–36). 'The Limits of Empiricism'. *Proceedings of the Aristotelian Society, New Series 36*: 131–150.

Rycroft, C. (1979). *The Innocence of Dreams*. New York: Pantheon.

Wittgenstein, L. (1921). *Tractatus Logico-Philosophicus*, trans. D. Pears and B. McGuinness. London: Routledge & Kegan Paul.

Wittgenstein, L. (1929). 'A Lecture on Ethics'. In *Ludwig Wittgenstein: Philosophical Occasions, 1912–1951*, eds J. Klagge and A. Norman. Indianapolis, IN/Cambridge: Hackett.

Wittgenstein, L. (1932–5). *Wittgenstein's Lectures: Cambridge, 1932–1935*, from the notes of Alice Ambrose and Margaret Macdonald, ed. Alice Ambrose. Chicago, IL: University of Chicago Press.

Wittgenstein, L. (1933–4). *The Blue and Brown Books*. Oxford: Blackwell.

Wittgenstein, L. (1937). 'Cause and Effect: Intuitive Awareness'. In *Ludwig Wittgenstein: Philosophical Occasions 1912–1951*, trans. R. Rhees, J. Klagge, and A. Norman (pp. 368–426. Indianapolis, IN/Cambridge: Hackett.

Wittgenstein, L. (1938). 'Lectures on Aesthetics'. In *Lectures and Conversations on Aesthetics, Psychology and Religious Belief*, ed. Cyril Barrett (pp. 1–36). Berkeley/Los Angeles, CA: University of California Press.

Wittgenstein, L. (1942–6). 'Conversations on Freud'. In *Lectures and Conversations on Aesthetics, Psychology and Religious Belief, 1938–1946*, compiled from notes taken by

Yorick Smythies, Rush Rhees, and James Taylor, ed. Cyril Barrett (pp. 41–52). Berkeley/ Los Angeles, CA: University of California Press.

Wittgenstein, L. (1945–9). *Philosophical Investigations*, trans. G. E. M. Anscombe, P. M. S. Hacker, and J. Schulte. Oxford: Wiley-Blackwell.

Wittgenstein, L. (1946–9). *Remarks on the Philosophy of Psychology*, 2 Vols: Vol. 1, eds G. E. M. Anscombe and G. H. von Wright; Vol. 2, eds G. H. von Wright and G. Lyman. Oxford: Blackwell.

Wittgenstein, L. (1977). *Culture and Value*, ed. G. H. von Wright, trans. Peter Winch of *Vermischte Bemerkungen*. Chicago, IL: University of Chicago Press.

Notes:

(1.) It is puzzling for Wittgenstein to raise the issue of infallibility at all, since nothing in Russell's argument in support of the view that we sometimes *experience* causality rather than inferring it inductively requires a discussion of infallibility. For like 'intuitive recognition', which does not appear to be Russell's term, 'immediate awareness' is ambiguous. If we are going to paraphrase Russell's claim in terms of 'immediate awareness', it's clear that what Russell meant was simply that when we experience an individual instance of causality, that experience is not mediated by any *inference*; in this respect, the experience of causality differs from inductive inferences, by which causal claims can also be derived.

(2.) The argument in [III] receives a very different critical treatment in Frank Cioffi's essay, 'Wittgenstein's Freud' (Cioffi 1969: 108–9); for a response to Cioffi, see Levy (1996): 45–56.

(3.) A glimpse of what Wittgenstein sought can be found in Charles Rycroft's *The Innocence of Dreams* (1979), wherein several ideas also to be found in Wittgenstein's Aesthetic Model are employed to support an alternative conception of psychoanalysis, as Jeremy Holmes and Edward Harcourt have noted (Holmes 2010: 183; Harcourt 2017: 661). For Wittgenstein, the Aesthetic Model was not only vital to his thinking about psychoanalysis; besides much else in his philosophy, some version of it probably also explains his low regard for Darwin's theory of evolution, which he thought lacked 'the necessary multiplicity', and so could not account for life's observable variety (Drury 1984: 160–1). In saying this, Wittgenstein need not be assumed to have been expressing hostility to the very idea of evolution itself. For until recently, biologists, unlike Darwin himself, very widely regarded the sexual selection that Darwin expounded (which rests on an aesthetic model), as *reducible* to natural selection ('survival of the fittest'). So Wittgenstein's dissatisfaction with Darwin can be seen as prophetic of the return to non-reductive, aesthetic evolution, alongside natural selection. Aesthetic evolution is argued for by Richard O. Prum (2017).

Wittgenstein and Psychoanalysis

Donald Levy

The Levy Group, Inc

Oxford Handbooks Online

Abstract and Keywords

Ricœur's reading of Freud is one of the most comprehensive, perceptive, and judicious explications of Freudianism—one that begins with his early 'Project' of 1895 and culminates with the last book that Freud published, Moses and Monotheism. Ricœur is successful in exposing some of the weaknesses in Freud, and even more importantly, in showing why there is a need to move beyond Freud. He also develops the significant idea of a dialectical relationship between a hermeneutics of suspicion and a restorative hermeneutics of meaning—and that they are integral to each other. Ricœur is successful in showing how, if one relentlessly pursues the logic of Freud's thinking, it leads beyond Freud. But, even though he gives some indications of how such dialectic is to be developed, this remains a task (an Aufgabe) that lies ahead.

Keywords: dialectic, Freud, psychoanalysis, hermeneutics, meaning

Richard J. Bernstein

Introduction

I want to begin with a few personal remarks about my encounters with Paul Ricœur—for they are relevant to the topic that I will be exploring. I first met Paul in the late 1970s. One of our first meetings was at a conference on Hannah Arendt in Paris—one of the earliest conferences on her work in France. At the time, there was barely any philosophic interest in Arendt but Ricœur and I shared an enthusiasm for her work. And we were both interested in her conception of action. Paul sometimes used my book *Praxis and Action* (Bernstein 1971) in the courses that he taught at the University of Chicago. But what really united us was his growing interest in and contribution to hermeneutics. In the 1970s and early 1980s there was a new interest in hermeneutics in America, in part due to the English translation of Gadamer's (1975) *Truth and Method*. But there were other contributing factors such as Habermas's (1971) appreciation of the role of hermeneutics

as a form of practical knowledge that he explores in *Knowledge and Human Interests*. And on the Anglo-American scene there was the appearance of Charles Taylor's (1971) classic essay 'Interpretation and the Sciences of Man'. Even Thomas Kuhn came to appreciate that *The Structure of Scientific Revolutions* (Kuhn 1962) had a deep affinity with hermeneutics. In 1977 in *The Essential Tension*, Kuhn writes:

> What I as a physicist had to discover for myself, most historians learn by example in the course of professional training. Consciously or not, they are all practitioners of the hermeneutic method. In my case, however the discovery of hermeneutics did more than make history seem consequential. Its most immediate and decisive effect was instead on my view of science.

And he adds:

> In my own case, for example, even the term 'hermeneutics,' to which I resorted briefly above, was no part of my vocabulary as recently as five years ago. Increasingly, (p. 204) I suspect that anyone who believes that history may have deep philosophical import will have to learn to bridge the longstanding divide between the Continental and English language philosophical traditions.

> (Kuhn 1977: xiii)

I don't think it is an exaggeration to suggest that during these decades there was a sea change taking place—that in the humanities, the social sciences, and even the natural sciences there was a new appreciation for hermeneutics—for the significance of understanding and interpretation of meaning and its importance for understanding our being in the world. And clearly one of the leaders in this new sensitivity to hermeneutics was Paul Ricœur.

Ricœur's Freud

By 1970, Ricœur had already published his monumental *Freud and Philosophy*, which he subtitled 'An Essay in Interpretation' based on the Terry lectures that he gave at Yale. The full title of these lectures is 'The Dwight Harrington Terry Foundation Lectures in Religion in the Light of Science and Philosophy'. The statement of the purpose of these lectures says that 'it is desired that a series of lectures be given by men eminent in their respective departments … to the end that the Christian spirit may be nurtured in the fullest light of the world's knowledge and that mankind may be helped to attain its highest possible welfare and happiness upon this earth' (Ricœur 1970: vi). Although Ricœur was known for his work on the philosophy of religion, it must have taken an act of intellectual courage to give a series of lectures on Freud—who was one of the most outspoken and vigorous atheistic thinkers of the twentieth century. *Freud and Philosophy* is a remarkable book. Written more than forty years ago, it stands as one of the most judicious, thorough, and comprehensive explications of Freud's work and development. Ricœur displays an intellectual virtue that is all too rare. In Book II, the main part of his

study, entitled 'Analytic: Reading of Freud', Ricœur not only shows his subtle understanding of Freud's entire corpus but he brackets his own philosophic views to explicate the complex stages of Freud's development. It is a model of interpretation that stays close to Freud's texts and is always illuminating. Freud presents a deep challenge to Ricœur's own philosophical convictions. I have always admired those thinkers who have the imagination and the hermeneutical generosity to 'take on' what is alien and antithetical to their own way of thinking. Ricœur exemplifies this. In one of his essays he writes:

> For someone trained in phenomenology, existential philosophy, linguistic or semiological methods, and the revival of Hegel studies, the encounter with psychoanalysis constitutes a considerable shock, for this discipline affects and questions anew not simply some particular theme within philosophical reflection but the philosophical project as a whole. The contemporary philosopher meets Freud on the same ground as Nietzsche and Marx. All three rise up before him as protagonists of suspicion who rip away masks and pose the novel problem of the lie of consciousness and (p. 205) consciousness as a lie. This problem cannot remain just one among many, for what all three generally and radically put into question is something that appears to any good phenomenologist as the field, foundation, and the very origin of any meaning at all: consciousness itself. What in one sense is a foundation must appear to us in a different sense as a prejudice, the prejudice of consciousness.

> (Ricœur 1974: 99)

Ricœur is alluding to his famous idea of a hermeneutics of suspicion. I will have more to say about this, but here I want to emphasize Ricœur's awareness that psychoanalysis poses a deep challenge to the foundation and starting point of any phenomenology when he speaks of the 'lie of consciousness and consciousness as a lie'. Consciousness turns out to be just as obscure as the unconscious. Consciousness is not a given, not a starting point, not a presupposition that cannot be questioned; it will turn out to be a task and an achievement. The questions that Ricœur poses are: 'What is the meaning of the unconscious for a being whose task is consciousness?'; 'What is consciousness as a task for a being who is somehow bound to those factors such as repetition and even regression, which the unconscious represents for the most part?' (Ricœur 1974: 108–9).

Energetics and Hermeneutics

To appreciate what is at stake in these questions and to appreciate what is involved in Ricœur's encounter with Freud, we must turn to the outlines of his 'reading' of Freud. Ricœur distinguishes two aspects or dimensions of Freud's approach—what he calls 'energetics' and 'hermeneutics'. To a great extent 'energetics' refers to Freud's earliest scientific and positivist leanings where he attempted to present an understanding of psychic life that was completely dependent on a 'quantitative apparatus'—one that Freud originally associated with the operation of the brain. Energetics is based on an 'economic'

interpretation of psychical life which is initially represented in Freud's 'Project' of 1895. Ricœur tells us that the ' "Project" of 1895 represents what could be called a non-hermeneutic state of the system. Indeed, the notion of the 'psychical apparatus' that dominates this essay appears to have no correlation with a work of deciphering and 'is based on a principle borrowed from physics—the constancy principle—and tends to be a quantitative treatment of energy'.[1] As Ricœur notes 'this recourse to the principle of constancy and the quantitative hypothesis is the aspect of Freudianism that most resists the reading I propose based on the correlation between energetics and hermeneutics, between connections of forces and relations of meanings'.[2] What Ricœur brilliantly shows is that a close reading of the 'Project' (Freud 1895) already indicates the tension in Freud's thinking—the need for interpretation even to make sense of the psychical apparatus. Consequently, even when Freud (before *The Interpretation of Dreams* in 1900) was most deeply influenced by the quantitative natural science of his time and (p. 206) hoped to provide a psychology limited to physical principles, we can already detect the hermeneutic tensions in the 'Project'—the beginning of Freud's sense of the need for interpretation of *meaning*. Ricœur sees the development of Freudian theory in terms of 'the gradual reduction of the notion of psychical apparatus—in the sense of "a machine which in a moment would run of itself"—to a topography in which space is no longer a place within the world of action but a scene of action where roles and masks enter into debate' (Ricœur 1970: 70). But even this characterization does not quite indicate what is distinctive about Ricœur's reading of Freud. Freud *never* gives up the model of energetics or economics—although it gets transformed in the course of his development of the topological models of the psyche—first, in terms of the unconscious, preconscious, and conscious, and later in terms of the id, ego, and superego. Even though the problem of deciphering and interpreting becomes more prominent, what Freud elaborates is a distinctive 'mixed discourse' of energetics and hermeneutics. Both dimensions are essential for psychoanalytic discourse—and they do not stand opposed to each other. Energetics (force) implies a hermeneutics (meaning) and a hermeneutics discloses energetics. So what initially presents itself as an *aporia*—the contradiction of energetics and hermeneutics—turns out ultimately to be Freud's great accomplishment. Ricœur declares that the whole problem of a Freudian epistemology may be centralized in a single question: 'How can the economic explanation be *involved* in an interpretation dealing with meanings; and conversely, how can interpretation be an *aspect* of the economic interpretation?' This is the question that Ricœur seeks to answer—and he does so by carefully pursuing three cycles in Freud's development. What emerges in the course of Ricœur's reading is an elaboration of what he calls Freud's 'semantics of desire'.

The Hermeneutics of Suspicion

Let me remind you of the structure of *Freud and Philosophy*. It consists of three parts or books: 'Problematic: The Placing of Freud', 'Analytic: Reading of Freud', and finally 'Dialectic: A Philosophic Interpretation of Freud'. The use of the terms 'Analytic' and 'Dialectic' might suggest that Ricœur is using Kant as a template for his reading and

interpretation of Freud, but we will see that it is really Hegel's (1807/2018) *Phenomenology of Spirit* that inspires Ricœur. In the first part, the placing of Freud, Ricœur reviews his own understanding of language and symbol. In this opening part he introduces a distinction between two different types of hermeneutics: interpretation as recollection of meaning and interpretation as exercise of suspicion. And the three great practitioners of the hermeneutics of suspicion are Marx, Nietzsche, and Freud—each, in a different way, seeks to expose the hidden or latent meaning of what is manifest; each challenges interpretations that fail to acknowledge what symbols *conceal*. This discussion of the conflicting types of interpretation sets the stage for Ricœur's reading of Freud. And this reading contains its own inner complexity because Ricœur shows how in the course of Freud's development the understanding of energetics and hermeneutics is transformed. (p. 207) Against the prevailing view that Freud began his psychoanalytic investigations with a focus exclusively on the analytic situation of curing individual patients and only turned to problems of culture in his later writings, Ricœur argues that the *interpretation* of culture was always evident in Freud's writings. He also argues that when Freud introduces the Death Instinct and pursues its consequences in the great struggle of Eros and Thanatos, there is a radical transformation in the psychoanalytic understanding of culture.

In the culminating section of *Freud and Philosophy*, 'Dialectic', Ricœur offers '*a*'—not '*the*'—philosophical interpretation of Freud. Although Ricœur makes it clear that his study is oriented to this final section, I actually think it is the least satisfactory. I want to explore some of the tensions and unresolved issues in his philosophical interpretation. But first consider Ricœur's description of his journey:

> In my introductory presentation of Freud I regarded him, along with Marx and Nietzsche, as one of the representatives of reductive and demystifying hermeneutics. In this view I was guided by a taste for extremes: I saw Freud as having a precise place in the hermeneutic debate, opposed to a nonreductive and restorative hermeneutics, and in league with other thinkers who wage a combat comparable to his. The whole movement of this book consists in a gradual readjusting of that initial position and of the panoramic view of the battlefield governing it. In the end it may seem that in this indecisive combat Freud is nowhere because he is everywhere. That impression is correct: The limits of psychoanalysis will finally have to be conceived not so much as a frontier beyond which exist other points of view, *but rather as the imaginary line of a front of investigation which constantly advances*, while the other points of view filter through the dividing line. In the beginning, Freud is one combatant among many; in the end, he shall have become the privileged witness of the total combat, for all the opposition will be carried over to him.

> (Ricœur 1970: 59–60)

Now there is something unsettling about this passage. It already indicates a certain hesitancy in how Ricœur understands Freud as a privileged witness to the total combat.

One might think that after the careful explication of Freud as an exemplar of the hermeneutics of suspicion, then one can (or should be able to) demarcate clearly the boundaries and limits of this hermeneutics of suspicion—and confront it with other more restorative hermeneutic interpretations. And there are many passages in Ricœur that suggest that this is his project. But things are not quite this straightforward. Indeed, the above passage suggests that there isn't a sharp boundary that delimits psychoanalysis but rather that there is an imaginary line of investigation which constantly advances. Phrased in a different way, Ricœur's reading of Freud is so powerful and comprehensive that it becomes increasingly difficult to 'contain' it, to assimilate and reconcile it with other types of restorative hermeneutics. There is a disparity and tension between what Ricœur claims to show and what he actually shows.

Let me illustrate this with an example that is fundamental for Ricœur's project. In his initial discussion of symbol, Ricœur reiterates a theme that had been prominent in his earlier reflections on language, the double meaning of symbol—the way in which symbols *reveal* and *conceal*—'the showing hiding of double meaning' (Ricœur 1970: 7). (p. 208) Ricœur emphasizes this double meaning in his phenomenology of religion and he notes a parallel between the role of symbols in religion and psychoanalysis:

> The problem of double meaning is not peculiar to psychoanalysis. It is known in the philosophy of religion in its constant encounter with those great cosmic symbols of earth, heaven, water, life, trees, and stones, and with strange narratives about the origin and end of things which are myths … What psychoanalysis encounters primarily as the distortion of elementary meanings connected with wishes and desires, the phenomenology of religion encounters primarily as a manifestation of a depth or, to use the word immediately, leaving for later a discussion of its content and validity, the revelation of the sacred.[3]

Ricœur, in his Analytic, his explication of psychoanalysis, rigorously excludes any reference to the phenomenology of religion—any reference to *the revelation of the sacred*. But he does return to this theme in his Dialectic—in his philosophic interpretation of Freud. And here he wants to acknowledge the contribution and challenge of Freud—as well as the limitations of his hermeneutics of suspicion. He acknowledges the force of Freud's psychoanalytic critique of religion. The psychoanalytic analysis of illusion, the comparison of religion with neurosis, the genetic account (in both the individual and culture) of the origin of religion in *Totem and Taboo* (Freud 1913) presents a powerful challenge to any believer. Ricœur tells us that Freud's critique is effective against superstition and naive conceptions of religion.[4] He notes that Freud draws an analogy between the three basic stages of the childhood condition—'neurotic phase, latency period, return of the repressed'—and religion, but Ricœur goes on to say that the meaning of the analogy remains and must remain indefinite:

> All that can be said is that man is capable of neurosis as he is capable of religion, and vice versa. The same causes—life's hardship, the triple suffering dealt the

individual by nature, his body, and other men—give rise to similar responses— neurotic ceremonials and religious ceremonials ... [5]

This may well be true, but it doesn't strike me as a forceful response to Freud—to the Freud that Ricœur has so effectively explicated. The point of naming the final section of his study 'Dialectic' is to indicate that there is a genuine dialectic between the 'merciless exercise of reductive hermeneutics' and a restorative hermeneutics that acknowledges what Ricœur calls the signs of 'the Wholly Other'. But although this is announced as his project, I don't see that he really *shows* this in any detail. Ricœur laments the cultural movement that seeks to objectify and reify the Wholly Other. He writes:

> It seems to me, however, that this cultural movement cannot and must not remain external to the restoration of the signs of the Wholly Other in their authentic function (p. 209) as sentinels of the horizon. Today we can no longer hear and read the signs of the approach of the Wholly Other except through the merciless exercise of reductive hermeneutics; such is our helplessness and perhaps our good fortune and joy. Faith in that region of the symbolic where the horizon function is constantly being reduced to the object function: thus arise idols, the religious figures of that same illusion which in metaphysics engenders the concepts of a supreme being, first as substance, absolute thought. An idol is the reification of the horizon into a thing, the fall of the sign into a supernatural and supracultural object.

> Thus there is a never ending task of distinguishing between the faith in religion— faith in the Wholly Other which draws near—and belief in the religious object, which becomes another object of our culture and thus a part of our own sphere.[6]

And he concludes this section with a dramatic claim: 'Thus the idols must die—so that symbols may live.'

What precisely is Ricœur telling us, and how does it bear on psychoanalysis as a hermeneutics of suspicion and the restorative hermeneutics of the sign of the Wholly Other? Ricœur believes that there is a pernicious cultural tendency to reify the Wholly Other—to make the sacred into some sort of divine *object*. Freudian psychoanalysis plays a crucial role of exposing and critiquing this tendency to idolatry. The hermeneutics of suspicion cannot be bracketed or put aside. It is the fire that purifies faith and keeps us from idolatry. In the final analysis, there are not *two* types of hermeneutics that can be neatly separated from each other; rather they are dialectically related. The hermeneutics of suspicion exemplified in Freud's multifaceted critique of religion *informs* a restorative hermeneutics of the signs of the Wholly Other. And a genuine hermeneutics of restorative meaning must pass through the fire of merciless suspicion. This is what Ricœur wants to show. And in many ways it is an attractive resolution of the conflict of interpretations. But, as I have already indicated, although this is what he *says*, I do not see that he has actually *shown* how this dialectic works. Ricœur wants to limit Freud's critique to a critique of idolatry—a cultural tendency that is always a seductive temptation, but he says that this critique does not touch the heart of a genuine faith in the signs of the Wholly

Other. I don't want to deny that this may be true, and I certainly don't want to affirm that psychoanalysis can or will ever 'refute' a true faith. But the hermeneutics of suspicion is not simply limited to a critique of idolatry. It also raises suspicions about a faith directed to the sacred symbols of the Wholly Other. If there is or can be a genuine *dialectical* reconciliation between 'the merciless exercise of reductive hermeneutics' and the restorative hermeneutics of 'the symbols of the Wholly Other', then this must be shown and not simply affirmed. I think it is to Ricœur's credit that in many places he indicates his awareness that he has outlined a *task* and not a completed project. He speaks of the 'never ending task' of distinguishing the faith of religion from belief in a religious object. More generally, he admits that the project of showing 'how other types of hermeneutics which are foreign to psychoanalysis, is waiting to be constructed'. (p. 210) This is 'currently the most urgent task of a philosophical anthropology'. He limits himself to pointing out 'certain border zones within that vast field'.[7]

Hegelian Dialectics

I find a similar tension in the way in which Ricœur seeks to reconcile what he calls the dialectic of archaeology and teleology. It is this dialectic that presumably is the key to his philosophical interpretation of Freud. 'Archaeology' is the term that Ricœur adopts to indicate Freud's analysis of the Subject—as it is understood from the perspective of reflective philosophy. Ricœur understands 'the Freudian metapsychology as an adventure of reflection; the dispossession of consciousness in its path, because the act of becoming conscious is its task'. And the result of this adventure is a 'wounded Cogito—a Cogito that posits itself but does not possess itself; a Cogito that sees its original truth only in and through the avowal of the inadequacy, illusion, and lying of actual consciousness'.[8] In the movement from force to language, from energetics to hermeneutics we discover the very emergence of the semantics of desire. But from a philosophical perspective, the problem is how can we reconcile the economic model that is so fundamental and persistent in Freud with reflection. Ricœur wants to show that there is a way of reconciling Freud's economic model with reflection—with the task of becoming conscious: 'Can we understand this archaeology within the framework of a philosophy of reflection?'[9] This is the fundamental question that Ricœur seeks to answer in his philosophic interpretation. An affirmative answer depends on showing the dialectic of archaeology and teleology. And the way that Ricœur sets about answering this question is at once bold and imaginative. In what might seem like a detour, he explores this development in Hegel's (1807/2018) *Phenomenology of Spirit*, focusing especially on the crucial stage when Hegel introduces the concepts of desire and life at the beginning of the section on self-consciousness. Ricœur makes it clear that he is not interested in a facile and absurd eclecticism of Freud and Hegel:

> Hegel and Freud each stand in a separate continent, and between one totality and another there can only be relations of homology. I will try to express one of these homologous relations by discovering in Freudianism a certain dialectic of archeology and teleology that is clearly evident in Hegel. The same connection is

in Freud, but in reverse order and proportion. Whereas Hegel links an explicit teleology of mind or spirit to an implicit archeology of life and desire, Freud links a thematized archeology of the unconscious to an unthematized teleology of the process of becoming conscious ... I seek to find in Freud an inverted image of Hegel, in order to discern, with the help of this schema certain dialectical features which, though (p. 211) obviously operative in analytic practice, have not found in the theory a complete systematic elaboration.[10]

Let me elucidate what Ricœur is claiming.[11] Hegel's great achievement in the *Phenomenology* is to trace and make concrete the complex enriching dialectical unfolding of spirit—spirit that initially is rooted in consciousness. And when the 'truth' of consciousness is revealed, we discover that it leads us to self-consciousness that presupposes desire and life. Hegel's dialectic of negativity gradually leads us to an ever-enriching concrete understanding of intersubjective recognition. Nothing like this is thematized in Freud. But Ricœur employs his rich discussion of Hegel to highlight those aspects of Freudianism where such a dialectic of the development of consciousness is implicit and unthematized. There is an implicit teleology (in this Hegelian sense) in Freudianism. Ricœur finds three indications of this in Freud. The first is in the 'operative concepts' that Freud employs (but does not thematize). In the analytic encounter between therapist and patient there is a structural homology with Hegel's depiction of the struggle for recognition. This is evident in the dynamics of transference where (ideally) analysis is completed 'with the attainment of the two consciousnesses, when the truth in the analyst has become the truth of sick consciousness. Then the patient is no longer alienated, no longer another: he has become a self, he has become himself.'[12] The second way in which we can discern the implicit teleology of Freudianism is in the Freudian concept of identification. Ricœur claims that the notion of identification is problematic for Freud. And here again Ricœur appeals to Hegel to illuminate a lacuna in Freud. Freud's texts suggest the 'beginning of a dialectic of desire, in which negation is placed at the very center of desire':

> The possibility is thus opened of rereading Freud's writings from the standpoint of the reduplication of consciousness. The rule of this rereading would be the oscillation between a dialectic and an economics, between a dialectic oriented toward the gradual emergence of self consciousness and an economics that explains the 'placements' and 'displacements' of desire through which this difficult emergence is effected.[13]

The final indication of the implicit teleology of Freudianism is the question of sublimation. Freud's comments about sublimation are suggestive, problematic, and sketchy.[14] Freud never really explicitly explains this 'instinctual vicissitude'. Ricœur reviews the various texts that discuss sublimation in order to show the incompleteness and inadequacies of Freud's reflections. But the topic of sublimation opens up the possibility of developing the teleology that is implicit in Freudian archaeology. 'Ultimately, (p. 212) through the more highly elaborated concepts of identification and idealization, the empty concept of sublimation refers us back to the operative unthematized concepts of the Freudian

economics.'[15] All of this can be summed up in 'the task of becoming conscious, which defines the finality of analysis'. For Ricœur this is what is at once implicit and entailed in the famous claim that 'Where id was, there ego shall be'.

What are we to make of this philosophical interpretation—this attempt by Ricœur to show a dialectic of archaeology and teleology in Freudianism? As always, Ricœur's reflections are illuminating and thought-provoking. In the spirit of immanent critique Ricœur exposes weaknesses in Freud. And he presents a powerful case for showing why Freudianism needs to be supplemented with a type of teleology that would be homologous with the dialectical story of the development of spirit that we find in Hegel's *Phenomenology*. But at best, what we have are notes and suggestions towards such a dialectic—not the dialectic itself. We may fully agree with Ricœur's critique: 'Ultimately, the task of becoming I, of becoming the ego, a task set within the economics of desire, is in principle irreducible to the economics.'[16] But this negative point only sets the task for showing in *detail* how starting with the Freudian semantics of desire we can progress to the development of self-consciousness—a self-consciousness that achieves its satisfaction in another self-consciousness. And although at times Ricœur's rhetoric suggests that he has actually concretely developed this dialectic, I am not sure he would disagree with me that he sketches a project and a task that is yet to be fulfilled.

I want to try to show this vividly by considering a concrete instance of two different types of interpretation that Ricœur develops. Ricœur frequently returns to Sophocles' *Oedipus Rex* for obvious reasons. This is the text that initially served as the basis for Freud's introduction of the Oedipus complex. Quoting *The Interpretation of Dreams* (Freud 1900), Ricœur tells us:

> Freud writes 'His destiny moves us only because it might have been ours—because the oracle laid the same curse on all of us before our birth as upon him'; and, later on, 'King Oedipus … merely shows us the fulfillment of our childhood wishes' *(ibid)*. Our pity and terror, the famous tragic *phobos* is merely an expression of the violence of our own repression before this manifestation of our impulses.[17]

Ricœur tells us that this reading is 'possible, illuminating and necessary, but there is a second possible reading which is not so much concerned with the drama of incest and patricide *which actually took* place as with the tragedy of truth'.[18] Ricœur then develops a subtle reading of the Oedipus story that points to 'Oedipus at Colonus'. The tragedy of Oedipus is that the king's pride must be broken through suffering. The tragedy is about the hubris of adult guilt, about a man who considers himself unaffected by the truth. Oedipus' zeal is the zeal of ignorance and is revealed in his outburst against Tiresias. (p. 213) 'Thus Tiresias and not Oedipus is the center from which the truth proceeds.' 'The connection between Oedipus' anger and the power of truth is the core of the real Oedipus tragedy.'[19] In pursuing this illuminating interpretation Ricœur brings out further revealing details of the tragedy of *Oedipus Rex*. Ricœur illustrates the two types of hermeneutics, 'one oriented toward the resurgence of archaic symbols and the other

toward the emergence of new symbols and ascending figures, all absorbed into the final stage, which, as in the *Phenomenology of Spirit*, is no longer a figure but knowledge'.[20]

Now Ricœur himself raises the very question we (his readers) want to raise: how are these two different interpretations related to each other? 'As long as we remain within the perspective of an opposition between the two, consciousness and unconsciousness will answer to two inverse interpretations progressive and regressive.'[21] Ricœur wants to avoid the Scylla of simply stating a stark abstract opposition and the Charybdis of a facile eclecticism. 'We cannot simply add Hegel and Freud and give to each a half of man.'[22] Once again, the problem is to show in detail that there really is a genuine dialectic between these two modes of interpretation. But just as we arrive at the denouement, at the point where Ricœur will show us the dialectical relation between these two different types of hermeneutic, he leaves us in suspense. And he concedes that he has not yet answered the question we want him to answer—and which he claims is the heart of the matter. He concludes his essay 'Consciousness and Unconsciousness' with the admission: 'Finally, we have left unanswered the question of the fundamental identity of these two hermeneutics—an identity which leads us to say that a phenomenology of spirit and an archeology of the unconscious speak *not of two halves of man but each one of the whole of man.*'[23]

Conclusion

In conclusion, I want to make it clear precisely what I am arguing. I think that Ricœur's 'Analytic', his reading of Freud is one of the most comprehensive, perceptive, and judicious explications of Freudianism, one that begins with his early 'Project' of 1895 and culminates with the last book that Freud published, *Moses and Monotheism* (Freud 1939). I also think that Ricœur is successful in exposing some of the weaknesses in Freud, and even more importantly, why we need to move beyond Freud. I am deeply sympathetic with his claim that there is a dialectical relationship between a hermeneutics of suspicion and a restorative hermeneutics of meaning—and that they are *integral* to each other. And I also think he is successful in showing how, if we relentlessly pursue the logic of Freud's thinking, we are led beyond Freud—we are led to a set of concerns about (p. 214) consciousness, self-consciousness, and intersubjectivity that Freudianism does not adequately illuminate. But although he gives us many indications of how such dialectic is to be developed, I do not think he has actually achieved this. This is still a task (an *Aufgabe*) that lies before us.[24]

References

Bernstein, R. (1971). *Praxis and Action*. Philadelphia, PA: University of Pennsylvania Press.

Freud, S. (1895). *Collected Papers*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 1. London: Hogarth Press.

Freud, S. (1900). *The Interpretation of Dreams*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 4. London: Hogarth Press.

Freud, S. (1913). *Totem and Taboo*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 13. London: Hogarth Press.

Freud, S. (1939). *Moses and Monotheism*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 23. London: Hogarth Press.

Gadamer, H.-G. (1975). *Truth and Method*. London: Sheed & Ward Ltd.

Habermas, J. (1971). *Knowledge and Human Interests*, trans. J. Shapiro. Boston, MA: Beacon Press.

Hegel, G. W. F. (1807/2018). *The Phenomenology of Spirit*, trans. T. Pinkard. Cambridge: Cambridge University Press.

Kuhn, T. (1962). *Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press.

Kuhn, T. (1977). *The Essential Tension*. Chicago, IL: University of Chicago Press.

Ricœur, P. (1970). *Freud and Philosophy: An Essay on Interpretation*, trans. Denis Savage. New Haven, CT: Yale University Press.

Ricœur, P. (1974). *The Conflict of Interpretations: Essays in Hermeneutics*, trans. W. Domingo et al., ed. D. Idhe. Evanston, IL: Northwestern University Press.

Taylor, C. (1971). 'Interpretation and the sciences of man'. *Review of Metaphysics 25*(1): 3–51.

Notes:

([1]) Ricœur (1970: 69). [Eds: See also Hopkins, this volume.]

([2]) Ricœur (1970:69–70).

([3]) Ricœur (1970:7).

([4]) [Eds: See Blass and Cottingham, this volume.]

([5]) Ricœur (1970:533).

([6]) Ricœur (1970:530–1).

([7]) Ricœur (1974:143).

([8]) Ricœur (1970:439).

(⁹) Ricœur (1970: 453).

(¹⁰) Ricœur (1970: 461–2).

(¹¹) [Eds: See also Macdonald, this volume.]

(¹²) Ricœur (1970:474).

(¹³) Ricœur (1970:483).

(¹⁴) [Eds: See also Gemes, this volume.]

(¹⁵) Ricœur (1970:492).

(¹⁶) Ricœur (1970).

(¹⁷) Ricœur (1974:115).

(¹⁸) Ricœur (1974).

(¹⁹) Ricœur (1974:116).

(²⁰) Ricœur (1974:117).

(²¹) Ricœur (1974:118).

(²²) Ricœur (1974).

(²³) Ricœur (1974: 120; emphasis added).

(²⁴) This chapter was originally printed in *Ricœur Studies 4*(1): 130–9 by the University Library System, University of Pittsburgh. Copyright © 2013 Richard J. Bernstein. It is reprinted here with permission.

Richard J. Bernstein

The New School for Social Research, New York, USA

Oxford Handbooks Online

Abstract and Keywords

This chapter begins with a consideration of transience and mourning as a way of thinking about what a failure of imagination might be. It then considers why the 'fundamental rule' of psychoanalysis is truly fundamental: not only for psychoanalytic method, but also for understanding why Freudian psychoanalysis is of philosophical significance. The fundamental rule opens up a new meaning of speaking one's mind. The chapter argues that the psychoanalytic method is a deployment of reason, properly understood. For it is that activity of mind through which reason comes to an understanding of the mental activities—the thinkings—of non-rational and unconscious parts of the human psyche. It thus illuminates Socrates' claim in the Republic that reason ought to take the lead in organizing human life because it has insight into the whole psyche. The chapter examines what it is for the unconscious to function as a fate.

Keywords: imagination, reason, method, mourning, psychoanalysis, fate, melancholia, fundamental rule

Jonathan Lear

On Transience

IN November 1915, Sigmund Freud published an article on the fleeting nature of all things—ourselves included. The paper, 'On Transience', records 'a summer walk through a smiling countryside in the company of a taciturn friend and of a young but already famous poet'. The countryside was smiling, but neither poet nor friend smiled back.

> The poet admired the beauty of the scene around us but felt no joy in it. He was disturbed by the thought that all this beauty was fated to extinction, that it would vanish when winter came, like all human beauty and all the beauty and splendour that men have created or may create. All that he would *otherwise have loved and*

> *admired* seemed to him to *be shorn of its worth by the transience which was its doom.*

(Freud 1915b: 305, my emphasis)

Freud thought the poet was right about the transience of the world. He rejected as wishful the idea that there must be some eternal ground lying behind the beautiful appearances. He nevertheless thought that something was going importantly wrong in the poet's reflective evaluation of the situation. The sense of transience, Freud thought, ought to *increase* the value and joy in our experience. But, what really impressed Freud about this conversation was that his arguments with his companions seemed to go nowhere. Reason did not really seem to engage. (p. 222)

> My failure [to convince them] led me to infer that some powerful emotional factor was at work which was disturbing to judgement, and I believed later that I had discovered what it was. What spoilt their enjoyment of beauty must have been *a revolt in their minds against mourning*. The idea that all this beauty was transient was giving these two sensitive minds *a foretaste of mourning* over its decease; and, since the mind instinctively recoils from anything that is painful, they felt their enjoyment of beauty interfered with by thoughts of its transience.

(Freud 1915b: 306, my emphasis)

In Freud's view, his companions' imaginations leapt forward to an imagined future in which everything that mattered was destroyed; and then that image was used to attack the capacity to take joy in present beauty.

But what is wrong with that? Usually we use the phrase 'failure of imagination' to mean the person is not being imaginative enough; but in this case Freud seems to think his companions were being too imaginative, and imaginative in the wrong sort of way. But what about this imagining makes it the wrong sort of way? It cannot simply be leaping ahead to the future in order to inform the present. A similar imaginative leap could be used to enhance the sense of joy in the light of beauty's transience—and, in Freud's view, this would be an altogether healthy imaginative move. Nor can their problem be simply that their imaginings brought them a wistful lack of joy in the world's beauty. The failure does not consist simply in a failure to produce pleasure—though pleasure and pain may be involved in explaining the psychodynamics. What makes this revolt a failure, in Freud's opinion, rather than just a movement of imagination is that it disturbed his companions' judgement. Basically, the movement of imagination is disturbing the proper function of reason in two related ways. First, his companions are thoughtfully and reflectively drawing the wrong conclusion about the meaning of transience. Second, they are unresponsive to the salient conditions that ought to be the basis for changing their minds.

Freud says that what is going wrong is 'a revolt in their minds against mourning'. It is not clear what this means. He is no doubt drawing on our familiar experiences of mourning,

and perhaps ideas are percolating that will emerge a couple years later in his essay 'Mourning and Melancholia'; but Freud could not simply have meant a passing psychological state. For, he is trying to account for an organized way of life. The young poet and the taciturn friend were not fleeing mourning a particular dead loved one. Rather, the 'revolt' showed up in a stable way of life: a resolute refusal to take joy in the world's beauty (ostensibly due to its transience). Freud suggests a dynamic psychological account of how this works: a 'foretaste' of mourning triggers a 'recoil' from painful feelings. Still, what he is trying to explain is not an isolated psychological response, but a stable, character-like formation: this is how the poet lives with the beauty of the world. From Freud's perspective, it is actually a character-*deformation*: the poet and his friend are not living well with respect to the world's beauty.

But then what would living well with transient beauty consist in? Freud says that our sense of the transience of the world's beauty ought to *increase* the joy we take in it (Freud 1915b: 305-306). Now if what is getting in the way of this response is a revolt (p. 223) against mourning, this would suggest that a healthy response—*joy in transient beauty*—would be an unfettered exercise of a healthy and developed capacity to mourn. This may at first seem like a strange conclusion. After all, don't we think of mourning as the response to the death of a loved one? To be sure, this is the paradigm; but Freud seems to be suggesting that we take this paradigm and think outwards from it. By way of analogy, consider Aristotle's virtue of courage. The paradigm is on the battlefield, risking one's own life to protect the *polis*, Athens. But then as we reflect on the myriad occasions in life that require us to take a stand, to take some personal risks for a higher cause, we see that this condition of the psyche, courage, may be deployed at almost any moment in life. Although Freud's line of reasoning is not explicit here, it seems to take a similar turn. If we consider the broad structure of human life it seems to be marked by vulnerability to loss and separation. Not only do our loved ones die, it can happen that we outgrow our friends or they outgrow us. When things are going well we leave our own childhoods behind, at least, up to a point. Our relations to teddy do not remain the same. At the heart of human development is a capacity to bid adieu. It is in the spirit of Freud's writing to treat this capacity as a capacity to mourn broadly construed.

We do not yet know the marks and features of mourning when it is used in this broad sense. We have thus far isolated it via paradigm cases and an invitation to expand our conception of it in certain directions. But it is clear that for Freud the capacity to mourn is an important aspect of living well with the transience of things. And since we ourselves are transient beings living in a world of transient beings, it would seem that the capacity to mourn (broadly understood) is a virtue or a human excellence in Aristotle's sense. For what it is for creatures like us to live well is to live well with transience, separation, and loss—and Freud is isolating a psychic capacity whose function is to do just that.

So if mourning broadly understood is a form of living well it seems as though we might have a standard by which we can judge imaginative life in terms of better and worse. Are there distinctive forms of imagining that characterize failures to mourn? Are there forms of imagining that distort one's view of the world and thus get in the way of one's attempt

to live a meaningful life? Does failure to mourn disfigure reason in its proper activities? It would seem so if it is the task of reason to distinguish appearance from reality about the most fundamental matters. And if so, is there anything to be done about this? In my opinion, one needs psychoanalytic insight to be able to address these questions. And one needs a philosophical outlook to grasp their deeper significance.

The Fundamental Rule

When I talk about the need for psychoanalytic insight I do not mean to refer to any particular dogma or theoretical claim. Rather, I am interested in insights that come to light from psychoanalytic activity according to its characteristic method. That method is structured by what Freud called the *fundamental rule* of psychoanalysis. The rule is simple (p. 224) to state: *say whatever it is that comes to mind without inhibition or censorship.*[1] Freud was right that the fundamental rule is fundamental to psychoanalysis. But it is fundamental to psychoanalysis because it establishes a special arena for reason and imagination to meet and interact.

The injunction to say *whatever* comes to mind without inhibition or censorship gives us one sense of what we might mean by the free flow of self-conscious thinking. But it is an unusual sense. The analytic situation as structured by the fundamental rule both facilitates and enjoins a certain spontaneity of mind. A person is encouraged to let her mind wander wherever it will go and simply to say out loud whatever comes next. This is obviously not the 'freedom' of rational thinking in the familiar sense, of thinking according to the constraints of logos. Indeed, with the fundamental rule, the normal constraints of logos are *ruled out* as a reason for not speaking. But the situation is more complex than just the encouragement of this form of spontaneity. It is not merely that the analysand is saying out loud what is coming to mind, but analysand and analyst are also listening together to what is being said and not said. The analysand in particular is apperceptively attending to the movements of her thought, but it is a special mode of apperception.[2] On the one hand, the attention being paid is informed by logos in the familiar sense that the person is attuned to what makes more or less sense; but, on the other hand, the person is freed of the normal responsibilities that usually attend apperceptive consciousness. There is no burden to explain or defend what 'I think'— certainly, not in a reasonable account. As a result, the analysand is freed to develop a capacity to attend to the workings of her own mind; a capacity that tends to be underdeveloped when one focuses on the flow of rational thought. The analytic situation then is not a moment in which irrationality is promoted for its own sake. It is more like a Sabbath moment in which the familiar constraints of rationality—the inhibition and censorship of 'silly' or 'bad' thoughts, the giving of reasons and so on—are encouraged to take a rest (Lear 2016). The overall aim is the development of one's rational capacity—in the sense of one's capacity to distinguish appearance from reality, both about the world and about oneself. In this sense the fundamental rule encourages a teleological suspension of the rational for the sake of the rational.

In effect, the analytic situation as structured by the fundamental rule sets up an intimate space for reason and the workings of the imagination to meet and interact in fine-grained detail. As the analysand lies down and attempts to 'free associate', the associations tend of their own accord to become looser, at least by the standards of ordinary conscious thought. It is not an accident that in such an environment, analysands tend to remember their dreams. Dream memories come to mind because the analytic situation becomes dreamy.

It is a surprising and remarkable empirical discovery that no one can follow the fundamental rule. As Freud said, 'there comes a time in every analysis when the patient (p. 225) disregards it' (Freud 1913: 135n). In my experience that time comes almost immediately and recurs throughout the analysis. This means that the fundamental rule gives occasion for analyst and analysand to attend together both to the mind's spontaneous wanderings *and* to the internal resistance to any such opportunity. Attending to both makes possible a very rich and textured understanding of a person's imaginative life (including fantasy) and of how that imaginative life intersects with her self-conscious understanding.

And so, while the fundamental rule does encourage the free flow of self-conscious understanding in the sense of a self-conscious sense of the mind's spontaneous wanderings, it also facilitates the freedom of self-consciousness in the deeper sense of the development of a capacity of mind. That is, psychoanalysis is the development of a psychic capacity to attend to the imaginative workings of one's own mind as well as to internal efforts to disrupt them, and then to thoughtfully and effectively intervene simply through the self-conscious thinking one is doing. One develops a capacity of mind to change the structure of one's mind, immediately and directly, through the very understanding that brings about the change. In the remainder of this chapter, I shall try to make this claim clear.

Unconscious Fate

We are in a position to formulate a thought that is not explicitly in Freud, though there are suggestions of it throughout his work. The thought is this: *even at the level of the unconscious our being is an issue for us*. Unbeknownst to us for the most part, our imaginations are actively thinking through what kind of creature we are and what kind of world we inhabit. Issues of transience and omnipotence, power and vulnerability have long been on our minds, but in a manner that we do not easily recognize. Freud showed that this imaginative thinking has a peculiar form: a form that justifies us in conceptualizing it as a form of thinking, albeit one not constrained by the normal demands of rationality. If we are to come to grips with how our imagination works, we need to understand how ontological debate is possible at the level of the unconscious.

Freud saw that repetition was a hallmark of neurotic modern life—and he took that to be a hallmark of modern life. People were caught up in unhappy-making routines, often stretching far back in life, with myriad variations. And modern society did not have any

particular interest in undoing such structures. He tried to think through what this could mean.

Let me give a clinical vignette. Ms M—a professional in a large corporation—came into analysis because of persistent feelings of anger, irritation, and frustration. She and her partner would fight regularly, and she would also feel frustrated and angry at work, with friends and family. Outwardly she was successful at her job, though she did get into unspoken feuds with colleagues. She would, as she put it, make 'passive-aggressive moves' in which she would struggle with people while pretending just to be doing her job. Similarly, with her partner: she was managing to stay in the relationship, but it was (p. 226) regularly threatened by struggles, unacknowledged and acknowledged. Typically, some imagined third party was involved. So, for example, Ms M's partner has an old girlfriend with whom he remained friends. This was an occasion for Ms M to form angry fantasies about that relationship. She would express her fury in small gestures—like leaving dirty dishes in the sink—that could be denied if challenged. And if a fight did break out, Ms M would assume the posture of the pure voice of objective reason, wondering out loud why her partner was getting so angry. This would enrage him further, and Ms M would secretly enjoy his frustration. There is much to be said about this case, but I want to consider the broad structure. Freud focused on repetition—and that is how it appears when one looks consciously at the phenomena—but he also had the genius to recognize that the unconscious is a mode of imaginative life that has its own form.[3]

One of the hallmarks of unconscious mental functioning is *timelessness*. It is important to understand what this means. There are of course remarks in Freud that suggest that memories of events from long ago persist in the unconscious and continue to have a causal effect. But there is a deeper meaning, a claim about the temporality of unconscious thinking. From the timeless perspective of the unconscious, there is no such thing as repetition. There is no same *again*. From an unconscious perspective there is just timeless perdurance: a tenseless *the same*. In this way, the unconscious functions like a fate.[4]

The idea that we experience our unconscious as fate is sometimes not that far from conscious awareness. People may enter analysis with an inchoate worry that they are somehow destined never to have an intimate relationship, for example; or, that they will never quite express the creativity of which they are capable. Freud showed that in an important sense they are right.

In the case of Ms M, the fate was *that life shall be frustrating*. This fate hung timelessly over her life—and uncannily informed the daily events of which Ms M was consciously, temporally aware. (In Aristotelian terms, it functioned like a formal and final cause.) Ms M was painfully, angrily aware of life's frustrations; yet unaware of how active she was in shaping her life to fit this fate. She would focus on real-life slights; but she also would construe ambiguous moments so as to trigger her sense of frustration. Occasionally she would surreptitiously provoke colleagues to stand in her way. In her imagination she would set up triangles: one person purportedly favouring another over Ms M. So on those

occasions when she was blocked in some way, she would feel defeated in a competition. She did not realize she was doing this. But having succeeded in arousing opposition, she would then be filled with rage and fantasies of revenge. It was as though Ms M lived inside a geodesic dome of frustration and anger that she mistook for the world.[5] (p. 227) The dome was made up of petite triangles—shaped around two others—made up of rivalry, frustration, anger, and revenge. They each looked like discrete problems, but they had a similar shape and together they made up a world. That is, although there was no conscious overall act of will, Ms M's mode of living otherwise took the form of a resolute refusal to take joy in the beauty of the world.

Freud also said of the unconscious that it was indifferent to contradiction.[6] The point is not that the unconscious allows us to believe contradictions because the beliefs reside in different parts of our mind.[7] Rather, our unconscious imaginings are basically not troubled by contradictory evidence. Ms M was adept at passing over things that were seemingly 'going right' in her life. And she was creative in interpreting these apparently good events as *mere* appearance. That is, she was adept at creating and maintaining a false appearance–reality distinction. To put it in philosophical terms, her imagination distorted her capacity for reason. Sometimes she would subvert an event that was going her way. For example, she would on occasion instigate a fight—all the while thinking it was her partner who spoiled things. The fight would excite her, though it left her feeling angry, frustrated, and alienated from her partner.

Now, the point about these unconscious fates is that they are akin to a philosophical claim about our finite condition. *That life shall be frustrating*: We are creatures who consciously recognize that we are not omnipotent, that we inhabit a world that to a significant extent exists independently of us, that we have desires and longings that we are unable to satisfy. From this perspective, fate seems like a claim about who we are and what is possible for us. It is as though the unconscious is working in a philosophical direction—what is it like to be the finite creatures we are?—only it lacks reason.

The irrationality consists not just in the unconscious' imperviousness to contradictory evidence, but in its unrelenting insistence. These two features of unconscious mental functioning—timelessness and indifference to contradiction—help establish an *imaginative* a priori. Ms M's imagination sustained a teleological structure that interpreted all of life's unfolding events in its terms: namely, *that life shall be frustrating*.

This imaginative a priori is further facilitated by forms of unconscious imaginative thinking Freud discovered and called primary process; in particular, displacement and condensation. In primary process almost any association can be made, based on similarities in sound, smell, touch, contiguity in time or space, or the merest contingency. This means that virtually anything can be drawn in to symbolize life's teleological structure as destined for frustration. If she dreams of a door, for example, it quickly becomes *a door she cannot go through*—'shut, shut out, shut in, alone'—the imaginary meaning of a life summed up in a seemingly benign image.

(p. 228) Melancholia

It might at first be surprising that this unconscious fate-like structure—that *life shall be frustrating*—is basically an ego strategy. It is a mode of living through which Ms M expressed her understanding of herself as a finite creature. Certainly, it came as a surprise to Freud—and it came relatively late in his career—that the unconscious could not be equated with the repressed. The ego regularly works in ways of which we are unaware—and to which we are deeply resistant to finding out. But we are not here dealing with repression and the repressed.

> . . . we find ourselves in an unforeseen situation. We have come upon something in the ego itself which is also unconscious, which behaves exactly like the repressed—that is, which produces powerful effects without itself being conscious and which requires special work before it can be made conscious … we must admit that the characteristic of being unconscious begins to lose significance for us. It becomes a quality which can have many meanings …

(Freud 1923: 15–18)

What I think Freud means by this last statement is that the characteristic of being *repressed* may no longer be the primary concern of those aspects of unconscious mental functioning that emerge for us in psychoanalysis. We need to grasp how these fate-like ego strategies work. For Ms M, the insistence on life's frustration had almost the opposite of its official meaning. It was a way of taking control. The world could not inflict frustration on her because, in fantasy, she would get there first, and inflict it on herself. By insisting upon frustration, Ms M protected a hidden sense of invulnerability. This came out in dreams and daydreams of being all-powerful: able, for instance, to rip out the heart of an opponent and devour it or shred her enemy with sabre-like fingernails. The frustration and anger that Ms M consciously felt covered over omnipotent fantasies of revenge. She also felt omnipotent when, in a fight, she took up the position of unemotional reason looking down objectively on her opponent: wondering out loud in an 'unemotional voice' why his problem would get him so upset.

Rather than focus on the content of Ms M's fantasies, consider their structure: one level of unconscious ego inflicts a fate *that life shall be frustrating*, and a level of unconscious fantasy wishfully expresses omnipotence. This broad-scale structure has many variants and I suspect it is not unusual in contemporary urban bourgeois life. It is an unhappy-making solution, but one compatible with managing the social and work demands of readers of the *New York Times* (or the *Guardian*). Within the psyche of the high-functioning neurotic there is in some sense an ongoing ontological conflict: one side insisting on the person's finite nature, the other claiming omnipotence. In Ms M's case, the conflict culminates in a standoff in which Ms M expresses both her sense of finiteness and her sense of omnipotence in an unhappy but triumphant compromise of frustration.

I want to characterize this overall psychic structure—unconscious claims of omnipotence set over against unconscious insistence on the fated disappointment of a (p. 229) finite creature—as *melancholic*. Not melancholic in the familiar sense of a particular psychological state, mood, or emotion, such as depression, though, of course, such emotions might be part of this melancholic structure. This 'melancholia' should be understood broadly as the unhealthy counterpart to the 'mourning' (see Freud 1917) I mentioned at the beginning. It is a psychic configuration that makes mourning impossible. I want to say that it is a stuck insistence that the world be, in this case, a frustrating place. But what does it mean to be stuck? It cannot mean simply that a person insistently brings everything back, again and again, to the same idea. Think of creative hedgehogs, like Plato, Dostoyevsky, or Proust.[8] Think of those joyous saintly figures who are able to bring each human experience back to the idea that humans are created in God's image; or of Kierkegaard's claim that purity of the heart is to will one thing; or of Bob Dylan's claim that all his songs end by wishing his listeners good luck.[9] Just because of repeated insistence, we do not thereby think of them as stuck. For the same reason, it is not sufficient for being stuck that one is living with an imaginative a priori that tends one to experience things in certain ways, tends one's reason down certain paths. Such imaginative direction may be internal to one's creative flourishing.

Ms M is stuck because she is stuck *in falsity*—and in two related ways. First, in Ms M's case reason has not merely been influenced by imagination, it has been subverted by it. Her capacity thoughtfully to distinguish appearance from reality has been distorted, so that she has become adept at interpreting any purported successes as mere appearance. She similarly interprets apparent failures immediately as a real-life basis for frustration. She is stuck in a false conception of reality. And any attempts at further reflection only feed through this short-circuited loop. Second, she is stuck in a false understanding of self and world because she is unknowingly entangled in a poor solution to psychic conflict. She repeatedly experiences the world as frustrating because, almost unbeknownst to her, her imagination is taken up with childish fantasies of omnipotence. These fantasies of course present significant challenges in adult life. If they are not worked through and modified through further imaginative work, they threaten to break through in demands and claims that disrupt one's position in the social world. They can drive one crazy. The 'solution' that Ms M hit upon—probably early in life—is to cover over those fantasies with an omnipotent insistence on life's frustrations. So the falsity here is a false understanding of oneself, and a false sense of where one's frustration is coming from.

This is, I think, what Freud means when he says that the neurotic is withdrawn from reality. Ms M cannot experience the variety of life's events—and thus their liveliness—because she has already interpreted them in the frame of frustration. All other beings are and must be frustrating beings. In this way, a largely unconscious ontological stance is held in place by a conflicted, but fairly stable psychological formation. And although Ms M may exercise imaginative creativity in enforcing this world view, it is wearying (p. 230)

always to be brought back to the *telos* of frustration. In this sense, melancholia, understood as an ontological stance, is a failure of imagination.

Transference

Freud came to see that these forms of imagining attempt to structure the entire analytic situation. He called this phenomenon transference, and said of it:

> . . . transference is itself only a piece of repetition, and the repetition is a transference of the forgotten past *not only on to the doctor but also on to all the other aspects of the current situation.*
>
> (Freud 1914: 151, my emphasis)

As is well known, Freud had an interest in the repetition of forgotten scenes from the past. But there is also in his work a strain of thinking that the significant repetition is not so much of a past scene, but the re-enactment of a past form of imaginative life. The purported past scene—whether or not it actually happened—functions as symbolic representation of this imaginative form of life. The best way to understand this, I think, is not in terms of repetition or reproduction, but as the self-maintaining imaginative activity of an unconscious teleological structure.

In his first serious reflection on transference—in the aftermath of his painful failed treatment with the patient known as Dora—Freud said that dealing with transference was 'by far the hardest part of the whole task':

> It is easy to learn how to interpret dreams, to extract from the patient's associations his unconscious thoughts and memories and to practice similar explanatory arts: for this the patient himself will always provide the text. Transference is the one thing the presence of which has to be detected almost without assistance and with only the slightest clues to go upon . . .
>
> (Freud 1905: 116–17)

There are, I think, two reasons Freud found dealing with transference so difficult. First, in transference there is a weird undertow in which the analyst is pulled by the currents of another person's imagination. One finds oneself in the midst of an intermingling of subjectivities that still needs to be better understood. And it is especially difficult to grasp what is happening if, as with Freud, one's self-image is of a doctor taking an objective stance towards a patient whose subjectivity is *over there*, firmly inside the patient. Second, it is emotionally demanding to be drawn into a melancholic imaginative structure. One is assigned a rigid role, and one can come to feel that if it were up to the analysand, one would be stuck there forever.

It is Freud's genius to recognize that the way to get unstuck is to learn how to play with the transference. Freud put it this way:

> The main instrument … for curbing the patient's compulsion to repeat and for turning it into a motive for remembering lies in the handling of the transference. We (p. 231) render the compulsion harmless, and indeed useful, by giving it the right to assert itself in a definite field. *We admit it into the transference as a playground in which it is allowed to expand in almost complete freedom* and in which it is expected to display to us everything in the way of pathogenic instincts that is hidden in the patient's mind. Provided only that the patient shows compliance enough to respect the necessary conditions of the analysis, we regularly succeed in giving all the symptoms of the illness a new transference meaning and in replacing his ordinary neurosis by a 'transference-neurosis' for which he can be cured by the therapeutic work. *The transference thus creates an intermediate region* between illness and real life through which the transition from one to the other is made.

(Freud 1914: 154, my emphasis)

I would like to take seriously the idea that working with the transference consists in analyst and analysand learning to play together. I will give a brief clinical vignette that illustrates what I mean by this, and then discuss why it matters. I hope to show how play is an imaginative activity of special philosophical and psychological significance.

After a few years of analysis, Ms M became more aware of her tendency to feel frustrated and angry with colleagues and friends, and with her partner. She was weary of it, but she also felt despondent about ever breaking out of the cycle. She still had angry feelings, she still had vengeful fantasies, but she was now able to notice them consciously in the moment, and keep them to herself rather than just act them out. In that way, Ms M felt her life was much less self-destructive than it had been: she was still in her relationship, for example, rather than having one more explosive break-up. But she continued to feel frustrated and angry.

In the transference, Ms M was regularly angry with me: for my silences, for what she experienced as my spare analytic comments (which I experienced as a friendly but definite analytic stance). She experienced me, at least at times, as withholding from her, and accused me of being a 'silent, passive-aggressive type, just like my mother'. This was not all she felt towards me, but these angry feelings of frustration would recur. It also made her angry when I pointed out that she was using our analytic relation to set up a situation of frustration. She would then turn her anger on herself: she must not know *at all* how to do analysis. She said she felt helpless and confused. In the transference I became a parent who would not come to her aid, and left her 'in the crib with dirty diapers'. And yet, there was *also* something unreal about all this for her. It was as though she was starting to wake up to the dream-like status of her waking life. This was a sign we were in the 'intermediate region' Freud talked about.

Ms M became ever more frustrated with me in my other roles as a professor and as a public speaker. She felt blocked from my classes and imagined I had a very good time with my students. She had read my books and was frustrated she couldn't learn from me

what I was thinking now. One day she mentioned that she had seen a notice that I would be giving a lecture in Chicago. She moved quickly on, to talking about a frustration in her work situation. It was as though bare mention was enough. I drew her attention to the brevity of her mention—and asked her if she might have felt the need to move away from it. She became angry with me for bringing it up. But she did say that if I hadn't said anything, she would have interpreted that as me tacitly giving her permission to go. Now (p. 232) that I had brought it up, that must mean I did not want to her to go. Either way I was giving her a hint about what to do.

This led to a wealth of associations from childhood. She had spent much of her youth sneaking around looking for clues, afraid to state her desire openly. In her pre-teens her mother had become worried about Ms M gaining weight and had put her on a diet. Ms M would sneak around for food. My not speaking much reminded her of not being fed. If she got caught—or if she refused to eat the 'disgusting' food her mother did give her—she would be sent to the basement as punishment. It is there that she remembers having her first violent fantasies of revenge. So, in the moment I was the fantasy mother—in the playground of the analytic situation. In asking her about her brief mention of my upcoming talk I had *caught her* sneaking around. And, by supposedly giving her hints that she shouldn't go I was forcing her on a diet, keeping her ravenously hungry, preventing her from being fed by me outside the confines of the analytic situation.

But though the feelings were strong, we could both feel we were in an intermediate zone, an arena in which the reality or unreality of events is held in a weird kind of abeyance. Ms M circled through many associations but came back to her experience of me getting in the way of her coming to my lecture. A voice welled up from deep inside her, a plaint: '*I have never heard you speak!*'

I said, 'Can you hear me now?' (My comment was earnest but as I hope you can see, also playful. My mind had associated to a well-known mobile phone ad in which the person repeatedly tests whether there is still a communicative link.) I do not want to exaggerate the efficacy of any single moment and I cannot say with absolute confidence what happened. But this seemed to me a significant moment. Ms M was at first taken aback, slightly shaken. She laughed, she sighed, she fell silent, her body relaxed. After a moment she said, 'Yes, I *can* hear you now'.

Ms M began to draw links between various meanings of *I have never heard you speak!* The official voice of the ego was the voice of frustration and complaint: because she was in analysis with me, she was prevented from going to hear me in a lecture or a classroom. But Ms M had been in analysis long enough that when she paid attention, she could hear another voice just below the surface: a voice of omnipotent, victorious triumph. *I have never heard you speak!* That is, 'I have absolute, unassailable right to go whether you like it or not!' But then there was also a voice of dawning recognition. *I have never heard you speak!* That is, 'I have spent this time hearing you in a false role, speaking indirectly to me passively denying me permission'. It was also a voice of mourning: 'I have spent these years in analysis so busy listening for clues whether you endorse what I am saying or

whether you deny me permission that I have only dimly heard your analytic voice—which you have been speaking all along'.

It would take a paper of its own to explain what happened in the analysis to make it possible at this later stage for Ms M to hear my comment 'Can you hear me now?' as something other than just another frustrating remark. That is all it would have been or could have been in the opening phase of the analysis. We should have to describe how working together Ms M and I created a 'playground' in which Ms M felt safe enough to play. In particular, she felt safe enough to 'play' with me as being someone other than a frustrating figure in her life.

(p. 233) This is crucial: playing is playing with ontology. It is a different way of being with the being of things. In his discussion of child's play, the psychoanalyst Donald Winnicott famously introduced the terms 'transitional object' and 'transitional phenomena' to describe a child's play with, say, a teddy bear. Play, according to Winnicott, provides a resting place from the standard tasks of distinguishing subjective from objective, inside from outside, me from not-me. In play the issue of whether teddy is part of me or not—whether, for instance, teddy simply knows what I am thinking or I simply know what he is thinking—is left in abeyance. It is the ontological status of the teddy that is played with. This kind of play is precious to our living; and Winnicott suggests that, with children, 'good enough' parents have an intuitive sense to protect the play space.

> *Of the transitional object* [that is, teddy] *it can be said that it is a matter of agreement between us and the baby that we will never ask the question: 'Did you conceive of this or was it presented to you from without?' The important point is that no decision on this point is expected. The question is not to be formulated.*

(Winnicott 1982: 12)

In a 'good-enough' environment children have freedom *to play* with the issues of their being and the being of things around them; for example: *Is teddy part of me?* Or: *Is teddy up to me?* But they are able to do so because, with the help of a parental, supporting environment, certain defining questions are left unasked. To ask them is to dissolve the play space.

The analytic situation is a recreation in adult life of a play space. Analysts are sometimes caricatured as not saying very much; but one benign way to understand this is in terms of a reluctance to ask the questions that would inevitably bring play to an end. As we have seen, Ms M was able to start playing with me as a figure in her life. In particular, were there other ways of hearing me say 'Can you hear me now?' It was the play that put her in a position where she could then effectively ask a question about the frustrating figure who would not let her go to the lecture. Was that figure really located *over there*, in me? Or was it a figment of her imagination? It was play in the analytic space that enabled her to raise imaginative possibilities—and then gingerly try them out: questions about the manner of her being and the manner of mine.

In learning to play, Ms M was building up a significant psychological capacity. For as she linked these different meanings together, she was doing more than playing with words. These meanings were voices inside Ms M's psyche expressing different outlooks, each of which was hers. By going back and forth between them she was in effect linking these voices together, bringing them into a different form of communication. This was an exercise in Ms M's developing capacity to integrate her psyche, precisely through her self-understanding that this was what she was doing. Ms M's playfulness was itself an integrating act.

This immediate and direct efficacy of psychoanalysis has not been sufficiently appreciated. To put it in psychoanalytic and philosophical terms: the ego can develop itself by developing a *practical* capacity with respect to its own functioning. This is a startling accomplishment: not just the capacity to change how one's mind functions, or even to change it on the basis of conscious understanding of how it functions, but to be able to (p. 234) change it immediately and directly through the efficacy of self-conscious thinking itself. On this occasion, when Ms M says, 'I *can* hear you now', she is actually creating the condition of hearing me. This is a remarkable power: the capacity for self-consciousness to take up and practically influence the workings of the whole psyche via its understanding of those workings. We need to understand much better than we do how this all works. In particular, what does influencing the workings *of the whole psyche* consist in? We typically see in such moments a move in the direction of wholeheartedness —as though the person's emotional life and will are coming together. But how does this work? It seems to me that we still lack an adequate answer.

Psychoanalysis as Human Reason

I believe we have reason to think of this psychic activity as a development and manifestation of human reason. It is practical reason applied to the mind's own workings. My inspiration for this thought is due to Socrates. In the *Republic*, he argues that it is appropriate for reason to rule the psyche because it is truly wise and exercises foresight on behalf of the whole soul.[10] Of course, we should be attentive to historical context, and not simply assume that Socrates is somehow speaking timeless, transcendent truths. However, there is a choice of historical contexts in which one might locate Socrates' dictum. I would like to suggest that we locate his remark not only in the local context of the ancient Athenian *polis*, but also in the broader historical context of human logos-making.

Socrates says it is appropriate for *reason* to rule. The Greek is *to logistikon*, which literally refers to our capacity to exercise and deploy logos. Our capacity for logos is a shared human capacity in which we attempt to react to reality by giving an account of it in language, an account we try to make adequate, one to which we hold ourselves and each other answerable, one which requires us to take responsibility even for the norms of adequacy. It is in terms of this logos-making capacity that we try to formulate a distinction between truth and appearance, both about the world and about ourselves. The

development of a capacity for logos-making is an historical development, but it has come to be internally linked with what it means to be human. That is, the capacity of logos-making has ontological implications. We can imagine circumstances in which our capacity for logos atrophied but sexual reproduction continued. I suppose one would say that the biological species survived, but there would be a crucial sense in which *we* human beings would go out of existence.

(p. 235) Now Socrates says it is appropriate for reason to rule, but we should not be constrained by a caricature of ruling. It is easy to have an image of reason as a benign, reasonable but firm monarch sitting in the throne room deciding which desires pounding at the gates shall be let across the drawbridge into the castle. But the Greek verb *archein* suggests more generally *taking the lead, functioning as a principle or starting point, lending form*. Socrates' claim, understood in its broadest context, is that our capacity for logos ought somehow to take the lead when it comes to the workings of the whole human psyche. In particular, our capacity for logos-making ought to take the lead with respect to the workings of the non-rational part of the soul. It ought to be able to lend creative direction to the imaginative workings of our soul. The appropriateness of *to logistikon* taking the lead depends on the claim that it is *wise* (*sophia*) and that it has *foresight* (*promethêian)* on behalf of the whole soul. That is, our capacity for logos-making has legitimacy only when it is actually moving in the direction of getting things right with respect to the whole soul.

I want to read Socrates and Plato as speaking beyond themselves. I see them as laying down constraints on what an adequate account of reason will turn out to be. To understand what reason is we need to figure out what the appropriate relations are between our logos-informed capacity to understand and the (non-logos-constrained) imaginative workings in our soul. What would it be for our capacity for logos to be wise and exercise foresight with respect to our imaginations? It seems to me that psychoanalysis is the attempt to work out what those appropriate relations are, in the therapeutic context of actually encouraging the development of those relations (not simply understanding theoretically what they should be). So understood, psychoanalysis is itself an exercise of human reason.[11] It is the activity of listening to the imaginative voices, the promptings and workings of the non-rational soul, and engaging them effectively with logos-informed understanding.

For Socrates in the *Republic* and Aristotle in the *Nicomachean Ethics* one of the central questions, if not the central question, was how reason, *to logistikon*, could exercise appropriate influence over the non-rational part of the soul. This is a question for which they could only provide a partial answer. I believe psychoanalysis ought to be understood as taking up that heritage and opening new paths for providing an answer.

Mourning the Loss of Omnipotence

It is, I think, an unusual insight, not one we could have anticipated in advance, that this power of self-consciousness to transform the mind requires the development of the capacity to mourn, in the sense that Freud used in his essay 'On Transience'. I said earlier that when Ms M entered analysis her overall psychic structure was melancholic in the (p. 236) sense that it was stuck, rigid, and life-flattening. It showed itself in a stuck form of living. Melancholic psychic structures are unhappy-making, but at the same time it is difficult to let go of them. The process that Freud called working-through is a mournful process. As we saw with Ms M, the moment of saying 'Yes, I *can* hear you' had something wistful about it, as though gaining the capacity to hear in less rigid ways was at the same time letting go of a familiar way of life.

Mourning is essentially an imaginative activity. In effect, we learn to play with death. In the paradigmatic case at the graveside, the death of a loved one is a time of grief. But in mourning our imaginations get busy. We are full of memories of the loved one, full of thoughts like, 'what would he say if he were here?', full of daydreams of how we will go on under the benign aura of his lasting influence. And we are full of playful fantasies—some conscious, some unconscious—in which he lives on inside us, in which we can hear him speak or he can hear us. For Freud, a healthy way of saying goodbye is learning how to play with our lost loved one. Mourning, broadly understood, is the playful way of leaving in abeyance the lostness of the lost one. It is in such play that we work out how to be as a transient being in the company of other transient beings.

In 'On Transience', Freud suggests we expand our conception of mourning—take it beyond the paradigm case of loss of a loved one and apply it more broadly to the structure of human life in general. One way of viewing Ms M's analysis is that in it she learned how to mourn her own omnipotence. That is, she learned how to play with its demise. And, because in this case we are dealing with structuring fantasies, learning how to play with the death of one's own sense of omnipotence is the same activity as allowing omnipotence to die. And as Ms M no longer has to deal with the threats of omnipotent fantasies breaking through, she can ease up on her need to cover them over with an unrelenting insistence on frustration. Mourning is the way we bid adieu to a melancholic psychic structure that has hitherto held us captive. In this way we again open ourselves up to life. For example, instead of forever facing false structures of frustration we open ourselves up to the genuine risks and corresponding joys of transient life. The question of how we are to be with our own unconscious is thus revealed as a question of how we are to be.

An Unusual Meal

Let me conclude with a clinical example that illustrates this point. About four months before the end of her analysis Ms M came in and reported a dream she had the night before.

I had to give a presentation. It has music in it. I wanted to make Thai food—that's somehow connected to the music. I'm excited about it. I have a recipe. I am going to go shopping for the ingredients. It will be spicy. Then later in the dream I'm anxious that I'm not prepared.

(p. 237) She comments that she is excited by the idea of making music and Thai food. She is not an expert, but is really interested in it. But she is anxious that as she goes shopping for ingredients she might not be able to find her way back. Then she asks:

Why *Thai food*? The word '*Thai*' could be significant, it has to do with *a connection* to music, but that is not clear.

She continues, 'I can't believe in the dream how much I'm enjoying my life'.

Then she says something that surprises me: 'I can see the Thai dish in front of me. But it is one someone else made. I have to get the things to make my own.' This is the first time I have heard about another person, so I ask her, 'There is someone else who has already done it?' She says yes, and then after a pause says 'I wonder whether you think that is you, and all I am going to do is copy you'. I say, 'What do you think?' She said, 'It feels like *my* idea. I can bring the Thai food into my presentation and *connect* it to the music. It feels right.'

This was the second time she said that she could *connect* Thai food to the music. So I said: 'You seem to think that *Thai* food would *tie* things to the music'. She responded, 'The Thai food is a tie—it could tie everything together—though I don't know how it will work out exactly. I have to go far away for the ingredients, but the food will be interesting, spicy, taste good, be unfamiliar to me.' She also wondered whether this would make a tie between the two of us, after the analysis ended. She would have to go off on her own to find her own ingredients and put them together.

I do not think it is too much of a stretch of the imagination to hear this as a dream-like comment on the end of the analysis. In her youth, she had been kept on a strict food regimen by her mother and this had led, in her own mind, to terrible feelings of frustration and anger; to sneaking and peeking; to hidden gluttony and purging; to getting caught and punished and having violent fantasies of revenge. In the dream, this structure is opened up. She is going to be able to make, eat, and present an exotic new dish—Thai food. It is not something with which she is familiar, but she is excited by the prospect. It feels fresh and new; a challenge. She has some anxiety that she might get lost as she looks for the spicy ingredients, but that does not dampen her sense of excitement.

And then through the loose associations of primary processes of the unconscious, *Thai* food is linked to *tie-food*. A heap of meanings are condensed into that image. In the dream, Ms M ties Thai food together with making music. She ties that together with her capacity to make a new kind of public presentation. The dream seems to be summing up

the meaning of the analysis for her in dream-like terms. The analysis has fed her. But it has fed her with a capacity to make her own tie-food. She can now make music for others.

We can see the truth of the dream demonstrated in the analytic session. As Ms M associates, she is able to tie her dream to her conscious thoughts about her dream. She is able to put her dream into logos-filled conscious thoughts that tie together in unusual ways the rational and non-rational parts of her soul. All of this—the dream and the later analysis of it—are playful and mournful anticipations of the end of the analysis. Ms M is (p. 238) anticipating the end, trying to sum it up as a way of getting ready to say goodbye. She is dealing with the transience of the analysis itself. Notice how lacking her goodbye is in the familiar structure of frustration and anger. This means that Ms M is not just mourning the end of the analysis, she is saying goodbye to a way of life that had for so long been her way of life. It is by thinking through her life in a psychoanalytic manner, interweaving reason and the imaginative workings of the non-rational parts of her soul, that she has been able to undo its false structure. This is the activity of reason taking appropriate lead of the whole soul. And note the imaginative hopefulness here for a future that cannot yet be comprehended. This open-ended hopefulness is integral to the willingness to give up hitherto imprisoning structures, however comfortable they once were. And it is a joy Ms M can now experience in the transience of life.[12]

References

Berlin, I. (2013). *The Hedgehog and the Fox: An Essay on Tolstoy's View of History*. Princeton, NJ: Princeton University Press.

Davidson, D. (1982). 'Paradoxes of irrationality'. In R. Wollheim and J. Hopkins (eds), *Philosophical Essays on Freud* (pp. 289–305). Cambridge: Cambridge University Press.

Engstrom, S. (2009). *The Form of Practical Knowledge: A Study of the Categorical Imperative*. Cambridge, MA: Harvard University Press.

Freud, S. (1900). *The Interpretation of Dreams*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vols 4–5. London: Hogarth Press.

Freud, S. (1905). 'Fragment of an analysis of a case of hysteria'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 7 (pp. 7–122). London: Hogarth Press.

Freud, S. (1910a). 'Five lectures on psycho-analysis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 11 (pp. 9–55). London: Hogarth Press.

Freud, S. (1910b). 'The future prospects of psycho-analytic therapy'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 11 (pp. 141–151). London: Hogarth Press.

Freud, S. (1912). 'Recommendations to physicians practicing psycho-analysis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 12 (pp. 111–120). London: Hogarth Press.

Freud, S. (1913). 'On the beginning of treatment'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 12 (pp. 123–144). London: Hogarth Press.

Freud, S. (1914). 'Remembering, repeating and working-through'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 12 (pp. 147–156). London: Hogarth Press.

(p. 239) Freud, S. (1915a). 'The unconscious'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 166–215). London: Hogarth Press.

Freud, S. (1915b). 'On transience'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 305–307). London: Hogarth Press.

Freud, S. (1917). 'Mourning and melancholia'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 243–258). London: Hogarth Press.

Freud, S. (1923). 'The ego and the id'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 19 (pp. 12–66). London: Hogarth Press.

Freud, S. (1925). 'An autobiographical study'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 20 (pp. 7–74). London: Hogarth Press.

Freud, S. (1969). 'Das Unbewußte'. In M. Bonaparte, A. Freud, E. Bibring, W. Hoffer, E. Kris, O. Isakower (eds), *Gesammelte Werke*, Vol. 10. Frankfurt: Fischer Verlag.

Lear, J. (2015). Freud, 2nd edn. New York: Routledge.

Lear, J. (2016). 'The Freudian sabbath'. In J. Kreines and R. Zuckert (eds), *Hegel on Philosophy in History* (pp. 230–247). Cambridge: Cambridge University Press.

Lear, J. (2017). *Wisdom Won From Illness: Essays in Psychoanalysis and Philosophy*. Cambridge, MA: Harvard University Press.

Mann, T. (1937). 'Freud and the future'. In *Essays of Three Decades*. New York: Alfred A. Knopf.

Pippin, R. (2014). *Hegel on Self-Consciousness: Desire and Death in the Phenomenology of Spirit*. Princeton, NJ: Princeton University Press.

Plato (2003). *Platonis: Rempublicam*, ed. S. R. Slings. Oxford: Clarendon Press.

Plato (2004). *Republic*, trans. C. D. C. Reeve. Indianapolis, IN: Hackett.

Rödl, S. (2007). *Self-Consciousness*. Cambridge, MA: Harvard University Press.

Winnicott, D. W. (1982). 'Transitional objects and transitional phenomena'. In *Playing and Reality* (pp. 1–25). New York: Routledge. (p. 240)

Notes:

(^{1}) See Freud 1913: 134–5; Freud 1900: 101–2; Freud 1910a: 31–2; Freud 1925: 40–1; Freud 1910b: 144; Freud 1912: 115. I discuss the importance of the fundamental rule in Lear 2017: 1–49.

(^{2}) I discuss this mode of apperception in Lear 2016. See also Pippin 2014.

(^{3}) 'To sum up: *exemption from mutual contradiction, primary process* (mobility of cathexes), *timelessness*, and *replacement of external by psychical reality*—these are the characteristics which we may expect to find in processes belonging to the system *Ucs*.' Freud 1915a: 187.

(^{4}) For a thought-provoking discussion of fate in psychoanalysis, see Mann 1937: 3–45.

(^{5}) See also the case of Ms A, in Lear 2017: 11–29. The broad-scale difference was that Ms A's world was one of disappointment and resignation; Ms M's was one of frustration and rage, passive aggression, and emotional explosions.

(^{6}) The German is *Widerspruchlosigkeit*. See Freud 1969: 286.

(^{7}) This is the route Davidson (1982) took and the problem he set out to solve. The solution is ingenious but I have argued (Lear 2015: 29–60) that it is not adequate to the psychological phenomena that show up in psychoanalysis.

(^{8}) See Berlin (2013) for an account of the use of this phrase.

(^{9}) 'Is there anything in addition to your songs that you want to say to people?

—Good luck.

—You don't say that in your songs.

—Oh, yes I do, every song tails off with 'Good Luck,—I hope you make it.'

('Bob Dylan Gives Press Conference in San Francisco, Part II', *Rolling Stone*, 20 January 1968).

(^{10}) Plato (2003, 2004), *Republic* 441e. I use the English words 'soul' and 'psyche' interchangeably as translations of the Greek word '*psyche*'. And thus by 'soul' I mean basically what Aristotle meant: that the soul is the principle of life in a living being.

(11) See Engstrom 2009; Rödl 2007.

(12) I would like to thank Matthew Boyle, Gabriel Lear, Anselm Mueller, Edna O'Shaughnessy, and the editors of this volume for comments on a previous draft; and Isabela Ferreira for copy-editing and comments.

Jonathan Lear

Department of Philosophy, University of Chicago, USA

Oxford Handbooks Online

Abstract and Keywords

This chapter surveys new developments in the theory of symbolism, primary process thinking, and dreams, and then revisits some of Freud's own material to suggest how his contribution has been variously neglected or misconstrued. Freud's broader treatment of symbolism, and his theory of drive as a motivation–cognition–affect matrix, suitably clarified, offer a rich and coherent context for understanding symbolization and symbolic activity across primary and secondary processes and along a pathological-normal continuum—from psychosis, dream, defence, and phantasy to healthy ego functioning, creativity, and waking rational thought. This material helps to bridge the supposed gap between Freud's metapsychology and his clinical theory.

Keywords: dreaming, drive theory, metapsychology, primary process, symbolism, symbolization

Agnes Petocz

> PSYCHO-analysis began as a method of treatment; but I did not want to commend it to you as a method of treatment, but on account of the truths it contains, on account of the information it gives us about what concerns human beings most of all—their own nature—and on account of the connections it discloses between the most different of their activities.

> Freud 1933: 156–7

It is tempting to regard the heterogeneity of the rapidly expanding contemporary psychoanalytic field as a sign of healthy progress. According to Christian (2017), the monolithic view of psychoanalysis as offering a unified theory of the mind 'no longer exists, and more recent authors have had to contend with a growing theoretical pluralism in psychoanalysis, as well as with findings from neuropsychology, cognitive science, empirical psychoanalytic research, and psychotherapy process and outcome studies' (2017: 127). Consequently, there is disagreement about the nature and extent of

Freud's contribution—where and how we are confirming his ideas, where and how we have disconfirmed them and successfully moved on, and where and how we are hampered by misguided allegiance to the classical Freudian past.

In my view the main problem for psychoanalytic theory today is to divorce fruitful development from fruitless fragmentation and the illusion of progress, and the main task is to show how Freud's unifying account of mind and behaviour, far from being challenged by empirical findings in other disciplines, is actually being confirmed as the only coherent theory that can properly accommodate them. My aim in this chapter is to address this theme in relation to three central interrelated topics: symbolism, the (p. 256) primary process, and dreams. In the first half of the chapter I provide a brief and perforce selective survey of contemporary work in these areas, with a view to identifying some of the major issues that occupy the attention of theorists and clinicians. In the second half I indicate what of Freud's material I think deserves revisiting in order to appreciate that his contribution to these issues has been neglected, and then offer some further thoughts towards clarification and synthesis.

Contemporary Developments: A Brief Survey

Freud (e.g. 1913, 1933) located his contribution within the domain of science and scientific psychology. Yet mainstream academic psychology—especially the Anglo-American tradition—has long dismissed his theory precisely on the grounds of its being an archaic relic of little scientific worth (but see Petocz 2004, 2015). This rejection continues even as psychology begins to embrace silently[1] many of Freud's ideas—in 'hot' and 'embodied' cognition, conceptual metaphor, attachment theory, evolutionary approaches to motivation and emotion, and various key factors in the dynamics of successful psychotherapeutic interventions.

In contrast, those within psychoanalysis concerned with promoting its scientific status have explicitly embraced developments in mainstream psychology's putatively scientific heartland, namely, cognitive science and neuroscience.[2] Both the information-processing model of the mind from cognitive science and the 'neuromania' (Tallis 2011) that has gripped psychology and extended to psychoanalysis in the new disciplines *neuropsychoanalysis* (Solms and Turnbull 2011) and *psychodynamic neuroscience* (Fotopoulou 2012; Fotopoulou, Pfaff, and Conway 2012) have clearly influenced psychoanalytic treatments of symbolism, the primary process, and dreaming.

Symbolism: Contemporary Developments and Questions

Freud singled out symbolism as a field 'in which we have new things to learn and do in fact discover new things every day' (1910: 142), as being 'perhaps the most remarkable (p. 257) chapter of the theory of dreams' (1916/17: 151), and as raising 'the most interesting and hitherto unsolved problems' (1923a: 242).

Symbolism, the Primary Process, and Dreams: Freud's Contribution

The received view is that throughout the history of psychoanalysis there have been significant changes in how the field has dealt with symbols, with one of the most fruitful being the evolving conception of symbolization (Levy 2012; Roussillon 2015). We have advanced beyond the narrow Freudian symbol as defensive substitute dominated by primary process thought and based on the unconscious as a container of repressed thoughts (see Deri 1984; Jones 1916). We have shifted from *symbolism* to *symbolization*, the latter encompassing two additional notions that are regarded as human developmental achievements especially important in clinical considerations.

The first is the idea—earlier promoted in the work of Klein, Sharpe, and Segal—that the capacity for *symbolization* is not merely a defensive and pathological process, as Freud maintained, but is also adaptive and progressive, allowing for healthy accommodation to the privations of life, as in the use of symbols in play, imagination, and creativity. Here, symbolization involves representing an absent object (rather than only a repressed one) (Rayner 1991), and can become the basis for later forms of 'adaptive regression' (Holt 2002; Kris 1952).

The second new notion—emerging from the work of British analysts such as Rycroft, Winnicott, and Bion—is that the capacity for symbolization is equivalent to, or the basis of, the capacity for thinking.[3] This is important to contemporary psychoanalysis since '*symbols are essential for thinking and for storing emotional experiences*' (da Rocha Barros and da Rocha Barros 2011: 879, my emphasis; see also Bucci 1997a, 1997b, 2008). In sum: '*Symbolization* is a broader concept than symbolism … it includes secondary process as well as primary process thought, and is considered *a process of linking and meaning-making*' (Aragno 1997; Freedman 1998). 'This process corresponds closely with Bion's (1962) concept of "alpha function", which he considered *a kind of mental digestion*' (Taylor 2010: 184, my emphasis).

Clinical applications have expanded accordingly onto the various ways in which symbolic processes can be prevented, damaged, or destroyed in trauma and psychosis (Levy 2012). Breakdowns in symbolization amount to breakdowns in the ability to engage in cognitive and emotional information processing. The aim now becomes one of the analyst using his or her containing function to restore and enhance the patient's damaged capacity for symbolization (Deri 1984; Green 1975). Taylor (2010) discusses several case studies, all of which turn on the analyst assisting the traumatized patient to slowly acquire 'the capacity to contain and modulate intense affects, and to use the symbolic system of language to think and talk about emotionally painful childhood experiences' (Taylor 2010: 195).

(p. 258) These developments prompt a number of questions. Is Freud's notion of symbolism indeed restricted to an archaic, primitive, regressive, defensive, pathological process, where the symbolized is invariably repressed? And is that notion to be distinguished from that of the use of symbolism as a progressive, healthy, creative, and functional ego process? If yes, what is the relationship between these two? Further, is Freudian symbolism to be distinguished from the later concept of symbolization as the basis of the capacity for thought and for the processing of emotional experience? If yes,

how does this new foundational notion of symbolization relate to the other? Did Freud overlook this last type of symbolization?

The Primary Process: Contemporary Developments and Questions

Freud pronounced the existence of primary process laws in the id to be 'one new fact, whose discovery we owe to psycho-analytic research' (1940: 164). He introduced the concept of the primary process as a type of mental functioning (unconscious and irrational) differing from secondary process (conscious and rational) thinking and occurring in phenomena (dreams, jokes, symptoms, etc.) that are not amenable to explanation in terms of the latter (Freud 1895, 1900, 1915a, 1940). He described the primary process both in terms of free-flowing (as opposed to bound) psychic energy involving mobility of drive cathexes, and producing, for example, the primitive, infantile hallucinatory attempts at gratification in dreaming (Freud 1895: 326–7) and in terms of the mental mechanisms of condensation and displacement that contribute to the special characteristics of the unconscious (Freud 1900, 1905: 88–9; see also Holt 1967, 2009; Rapaport 1951). The primary process operates in accordance with the pleasure principle (the seeking of immediate and direct gratification), but then becomes modified and constrained by the reality principle to produce the secondary processes that allow for the realistic pursuit of gratification (Freud 1911).

This contribution, pronounced revolutionary and original (Jones 1953; Robbins 2004), was initially faced with the same treatment as was the theory of symbolism. Gammelgaard (2006) remarks that early theorists were:

> working in a tradition of classical metapsychology trying to rescue the concept of primary process from its connotation of a primitive, chaotic form of thinking, making it useful for understanding sophisticated and creative thinking, [whereas] today there is a radically different approach to the concept of primary process and the whole of Freudian metapsychology.

(Gammelgaard 2006: 94)[4]

(p. 259) This radically different approach saw the concept of the primary process become absorbed into cognitive science (e.g. Bucci 1997a, 1997b). Gammelgaard (2006) discusses Bucci's echoing of Schafer's (1976) 'two languages' view, where the primary process is presented in terms *either* of biology (specifically nervous system energy) *or* of psychological systems in the mental apparatus. Bucci's preferred solution is to reject the concepts of energy and instinct and other central concepts of the metapsychology.[5] According to Gammelgaard, this rejection, arising from a failure to see *how* the physiological drive 'level' fits with that of the psychological and clinical, is the central problem of the fate of the primary process when it has been absorbed into cognitive science.

Nevertheless, more recent accounts have kept up with the gradual shift to second-generation (embodied) cognitive science (Kövecses 2005) and then to cognitive neuroscience. For example, Brakel and her colleagues (2002) originally provided an exclusively perceptual-cognitive view of the primary process as involving automatic, unconscious, typically associative, relatively fast, processing. This primary process, shared with other animals, is developmentally and evolutionarily adaptive, and its formal aspect can be indexed in humans via categorization of similarity based on attributes (e.g. shared individual features) as opposed to relations (e.g. shared global configurations), as assessed by the Geometric Categorisation Task (*GeoCat*) (Brakel et al. 2000). Brakel and colleagues initially claimed that the primary process is not per se emotional and pathological, but that—as demonstrated in their empirical research—people will naturally engage in it under certain emotional or pathological conditions (Brakel 2004; Brakel and Shevrin 2005). More recently, in keeping with the shift towards embodiment in cognitive science, they have conceded that the primary process is emotion-based (Cutler and Brakel 2014), and also that their account fits well with recent neuroscientific accounts (Bazan et al. 2013). They conclude that the primary process is not restricted to impulsivity, regression, and symptom formation, but is also healthy and adaptive.

Hopkins (2012, 2015, 2016, and this volume), in demonstrating the consilience of psychoanalysis and neuroscience, gives equal attention to biological and psychological aspects of the primary process. He argues that the concepts of variational free energy (FE) and the Bayesian brain developed by Karl Friston and colleagues (Carhart-Harris and Friston 2010; Friston 2010; Hobson et al. 2014) can be mapped onto Freud's early formulations in his *Project for a Scientific Psychology*. One suggestion of contemporary neuroscience is that the brain is a Helmholtzian inference-making machine, innately supplied with a 'virtual reality generator' that 'produces the fictive prior beliefs that Freud described as the primary process' (Hopkins 2016: 1). These prior beliefs undergo a process of Bayesian updating via inference. The brain has evolved to minimize FE via increasing accuracy and decreasing complexity. In waking both accuracy and complexity (p. 260) are increased, but in dreaming (and in symptoms) complexity, which is 'conceptually linked with emotional conflict and trauma' (2016: 8), is reduced. Mental disorder 'is thus caused by computational complexity' (2016: 1). This complexity theory 'integrates the full range of psychoanalytic hypotheses about dreaming, development, and disorder with the perspective of FE neuroscience, where they might receive additional testing' (2016: 10).[6]

With this line in mind, Modell (2014) claims that Freud's distinction between primary and secondary processes is fundamentally biological and one of the few scientifically confirmed psychoanalytic ideas, because neuroscientists have 'found the neural correlates of Freud's primary and secondary process' (2014: 813). He argues that Freud failed to see that there were *two* types of primary process: dreaming, which is wish-fulfilling; and waking, which functions instead as an inference-making tool useful for survival. Therefore, Freud was wrong to assign anticipation of danger to the secondary process and to claim that the primary process ignores reality.[7]

Here, again, questions arise. Is the primary process a biological process, a psychological process, or both? Is it restricted to what is archaic, primitive, and infantile, or is it also involved in creative and healthy functioning? Is it restricted to unconscious thinking in general? Was Freud wrong to claim that the primary process is 'cut off from relations with the external world' when in fact it is 'immediately and directly involved with perception and with an individual's affective communication' (Dorpat 2001: 449)? Is the primary process the initial condition of mind, preceding the secondary process in both evolutionary and developmental terms, or do both emerge together during development (Holt 2002, 2009)? Is the primary process evolutionarily maladaptive or adaptive? Is dreaming primary process different from waking primary process? Did Freud mistakenly restrict the primary process to the dreaming model?

Dreams: Contemporary Developments and Questions

According to Freud, dreams are the productions of the primary-process dream-work mechanisms that 'make use of any symbolizations that are already present in unconscious thinking' (1900: 349). They are the mental phenomena whose interpretation is both 'the royal road to knowledge of the unconscious mind' (1900: 608) and 'the most complete piece of work the young science has done up to the present' (1912: 265).

Contemporary developments have focused on three closely interrelated themes: dreaming as regressive, disguised wish-fulfilment versus dreaming as progressive snap shot of the mind; dreaming as basic thinking involving fruitful psychological work, a process that continues during the waking state; and dreaming as neurophysiologically based complexity reduction.

The first theme continues Jung's original objection to Freud's seeing the dream as vehicle of disguised repressed wish-fulfilment. In his revisiting of Freudian dream theory, Robbins (2004) points to a number of post-Freudian theorists—among them Fairbairn, Kohut, and Hartmann—who 'elaborated the idea that the dream is a snapshot dramatization or depiction, a kind of map of what is on the mind of the dreamer at a particular moment during the night, *as contrasted to an unconscious disguise*' (Robbins 2004: 357, my emphasis). This is related to the debate about whether dream bizarreness is the natural form of dreaming content or whether it is produced by the Freudian 'censor's' distorting activity (Boag 2006, 2012; Maze and Henry 1996; Petocz 2006). Robbins argues that Freud's material leads to two radically different and incompatible theories of dreaming—a primary process theory and a secondary process theory—and that the latter is favoured by Freud, by virtue of his references to representation and symbolism. Robbins complains that Freud neglected his own more revolutionary hypothesis that dreaming represents a 'qualitatively distinct protolinguistic expression of mind' (Robbins 2004: 380) involving hallucinatory actualization. In support of this he alludes to the Solms-Hobson dream debate, and to Braun's (1999) PET scan brain studies that indicate relative neural quiescence during dreaming of those parts of the brain that are involved in secondary process activity.

The second theme—arising from Bion's work—is that of dreaming as the basic form of unconscious emotional thinking. We are constantly dreaming, even when awake, and this is exactly what is prevented or damaged in psychosis. Schneider (2010) sees this as a great paradigm shift in psychoanalytic history:

> Freud does not believe that, by dreaming, the dreamer does psychological work that helps deal with unsettling emotional experience; rather, he believes that dreaming dissipates the threat of overwhelming anxiety caused by the tension of repressed impulses originating in childhood … Dreamwork accomplishes censorship, editing and compromising … There is no psychological work (yielding motivational advance) inherent in the disguising and smuggling function of dream-work.
>
> (Schneider 2010: 524)

According to the new view dreams are a manifestation of the symbolic function, a first step in the process of thought. This also challenges Freud's (1940) claim that psychosis and dreams are parallel because of the prevalence in both of the primary process. Freudian dream theory is not directly applicable to psychosis, because the psychotic dreamer's lack of associations and the concreteness of the hallucination preclude the usual psychoanalytic approach to their interpretation; they are not genuine dreams because they are not the result of dream-work mechanisms (Capozzi and De Masi 2001). Vinocur Fischbein and Miramón (2015) see Bion's work less as a paradigm shift and more as a special extension of Freud's which, based on clinical work with psychotics, (p. 262) draws out the implications for dreaming of the new notion of symbolization. Instead of symbolization being an expression of repressed material, Bion 'inverts the Freudian assertion … in his opinion it is symbolization and the dream-work-alpha which makes remembering possible' (Vinocur Fischbein and Miramón 2015: 987).

The third theme concerns the neuroscience of dreaming. This topic has been controversial largely by virtue of the long-standing Solms-Hobson dispute.[8] Hobson initially proposed an 'activation-synthesis' model of dreaming (Hobson and McCarley 1977) which rejected the Freudian wish-fulfilment theory, until Solms showed (Solms and Turnbull 2002) that dreaming involves activation of the same (i.e. dopaminergic) neural system as is involved in waking goal-seeking and reward-consummatory behaviour. Hobson insists that Freud's seminal dream book 'still exerts an antiscientific force on our understanding of dreaming' (Hobson 2014: 4). Yet, as Solms (2013) points out, Hobson's views (most recently with his shift to a 'protoconsciousness theory' (Hobson 2009)) are getting ever closer to Freud's. Indeed, his recent collaborative paper 'Virtual reality and consciousness inference in dreaming' (Hobson et al. 2014) has been drawn upon in Hopkins's (2016) complexity theory of dreaming and mental disorder (see also Hopkins, this volume).

Once more, questions arise. Is the contrast between dreams as disguised wishes and dreams as meaningful expressions or snapshots of the dreamer's mind a genuine contrast? Did Freud offer two radically different and conceptually incompatible theories

of dreaming, a primary-process theory of dreams as hallucinatory wish-fulfilments and a secondary-process theory of dreams as employing dream-work mechanisms that require symbolization? Are dreams equivalent to unconscious processing of experience, such that we can be said to be dreaming all the time? Or are they a particular type of unconscious processing that requires the sensory attenuation of the sleeping state, as in the complexity-reduction thesis? Is dreaming in psychosis different from normal dreaming, and, if so, how?

Obviously, these and the earlier questions are a mere subset of possible questions that emerge from the contemporary scene. There is insufficient space here to address them in the systematic or thorough way that is needed. So what I propose to do in the space that remains is first to indicate where and how revisiting some of Freud's material might help to clarify the issues and lead us to reassess his contribution, and then to offer some further thoughts towards clarification and synthesis.

Revisiting Freud's Material

The first step is to revisit Freud's theory of symbolism, showing how much broader it is than is usually claimed. This involves also addressing two problems in his treatment (the 'system unconscious' and the role of language) whose resolution has direct implications (p. 263) for thinking about symbolism, the primary process, and dreaming. From this the focus shifts away from the distinction between primary and secondary processes onto the distinction between what is motivationally primary and what is motivationally derivative, and this takes us to Freud's much-disputed drive theory of motivation.

The Freudian Broad (FB) Theory of Symbolism

Within Freud's thought may be discerned a general, unified theory of symbolism and symbolic activity—the Freudian Broad (FB) theory—which is both more encompassing and more sound than what has typically been identified, even by Freud, as his theory—the Freudian Narrow (FN) theory (Petocz 1999, 2011a, 2011b).

The FB theory understands something's being a symbol as involving not just individual objects (a church spire symbolizes the phallus), but also more complex actions and events (playing sport symbolizes—among many other things—Oedipal competition), in a process that can be located anywhere along dimensions of normal–pathological, conscious– unconscious, and conventional–nonconventional, although it is the controversial nonconventional symbolism that is the focus of Freud's attention. This theory supports the continuity of the symbolic function through so-called primary and secondary processes. However, the theory requires for its conceptual coherence and consolidation the rejection of two problematic propositions advanced by Freud, namely, that symbolism is the natural mode of expression of the 'system unconscious' (Freud 1912: 260; 1915c) and that symbolism is a language (1916–17: 166). Failure to appreciate these problems

has meant that the propositions have become absorbed uncritically into post-Freudian developments and influenced assessments of Freud's contribution.

However, Freud also offered an 'epistemic' conception of the unconscious (Freud 1912, 1915a, 1923b, 1933); what makes a mental state unconscious is simply that the subject is unaware of it. This coheres with Freud's preferred 'dynamic' conception of repression— that what is repressed is rendered unconscious by a competing motivational force (Freud 1915b)—as opposed to the 'structural' conception of repression—that the qualitatively distinct characteristics of the unconscious preclude entry into consciousness. Combining the epistemic unconscious with dynamic repression clarifies the distinction between primary and secondary processes as they apply to thinking in general and to symbolism in particular. First, the primary process as irrational and involving mechanisms of condensation and displacement is not restricted to unconscious processes, so this attempt —as also various later attempts—to characterize unconscious processes as differing *qualitatively* from conscious processes runs into difficulties (see Petocz 1999: 154–61).[9] Second, the primary process conceived as primitive infantile hallucinatory gratification is primary neither logically nor temporally, for the infant must first experience the reality that is subsequently wishfully hallucinated. This challenges (p. 264) the forced gulf between the primary-process id and the secondary-process ego. Thus, the worry about whether symbolism is a strictly primary process phenomenon, a characteristic of archaic thinking, or whether it is a secondary process, adopted by the ego either for purposes of disguise or as part of healthy ego development and necessary for sublimation, is misguided—symbolism occurs across the board, beginning with the infant's interested perceiving of similarities, resulting in associations (e.g. breast associations to round, soft objects) that can be revived at any later stage.

Further, Freud's failure to adhere consistently to the distinction between the conventional symbolism of language and nonconventional symbolism encouraged psychoanalytic theorizing not only to conflate the two, but also to go further, into conflating language and thought.[10] Nonconventional symbolism in Freud's sense is not a communicative system of arbitrarily or freely chosen signifiers lacking any similarities to their referents. It is a special case of nonverbal *motivated mistaken identity*, typically based on perceived similarities of form or function, when one thing (which an 'interpreter' subsequently identifies as the 'symbol') *is treated as if it were* or is *acted towards as if it were* another thing (which an 'interpreter' subsequently identifies as the 'symbolized'). It includes the person's having two simultaneous beliefs, whose contents (where, say, X is the symbol and Y is the symbolized) are 'Here is Y' alongside 'Here is X'. Moreover, in this process of *compromise formation* neither repression nor active searching for a substitute is necessary. The drive takes the substitute when the primary object is *unavailable*, which may involve mere physical absence. This process can be understood as an extension of the straightforward satisfaction of an unopposed drive impulse, but now involving two or more impulses/wishes satisfied within existing constraints.[11] Nevertheless, that symbolism is a type of compromise formation means that it involves some kind of mental division,[12] one part mistaking the symbol for the symbolized, while another part treats

the symbol non-symbolically—though there is variability in the extent to which conscious and unconscious processes are balanced.

In sum, the basis of (nonconventional) symbolism lies in four empirical facts: (i) the initial primary objects and consummatory activities of the innate instinctual drives; (ii) the long period of infantile dependence; (iii) the connection between the drives and cognitive structures, which leads to the 'interested' perceiving of similarities between (p. 265) the primary objects and other, non-primary, objects; and (iv) the unavailability (to some part of the mind), mainly through repression, of those primary objects, and the inhibition, mainly through repression, of the expression of the consummatory activities with respect to particular primary objects. The crucial point is (iii), for attention now moves from the distinction between primary and secondary processes onto the more salient distinction between what is motivationally primary and what is motivationally derivative or secondary. But further elucidation is required. Freud's approach alone respects the logical constraints and meets the psychological requirements that any theory of symbolism must follow (see Petocz 1999: 239–65, summarized in Figure 3, 241). One of those logical constraints is that a theory of symbolism, by virtue of symbolism's being a three-term relation in which one of the terms is a cognizing organism, must be a psychological theory.[13] There follows a set of psychological requirements, amongst which the most important is that the theory offer a *motivational* account to explain the ontogenesis of symbolism, the selection of the symbolized, and the unidirectional nature of the symbolic equation. The account must also explain mental plurality and the notion of 'interested' perceiving, and it must address a theme that, as shown in the first half of this chapter, crops up regularly in contemporary developments—the question of the relation between the biological and the psychological. For that, we need to revisit Freudian drive theory.

Freudian Drive Theory: A Motivation-Cognition-Affect Matrix

Freudian drive theory has long been controversial.[14] Even within Freud-friendly neuropsychoanalysis and psychodynamic neuroscience, there is the view (e.g. Watt 2012) that Freud's theory is outdated and clumsily mechanistic, ignores the Darwinian insight that homeostatic needs must have access to the perceptual and cognitive machinery of consciousness, cannot accommodate the primacy of attachment and other social needs, and has been superseded by the idea of multiple motivational/affective systems (e.g. Panksepp 1998). These criticisms could legitimately be applied to traditional drive theories—McDougallian, Hullian, Skinnerian—which are indeed inadequate. But *Freud's* conception of drives is very different; it goes far beyond any traditional conception, offering a motivational model that stands out amongst drive theories in providing the required complexity and sophistication. If, then, we come to Freud's material with our traditional conception of drive, it is difficult to appreciate Freud's originality, easy to see why 'the drives have proven to be a huge problem in psychoanalytic theorizing' (Solms and Zellner 2012: 51), and easy to reject prematurely (not via straw-manning but via *ignoratio elenchi*) Freud's formulations.

Symbolism, the Primary Process, and Dreams: Freud's Contribution

Freudian drive theory offers a realist and determinist motivational approach whose sophistication as a motivation-cognition-affect[15] matrix that is needed to explain all behaviour has not been fully appreciated. The seminal exposition of this argument appears in John Maze's (1983) book *The Meaning of Behaviour* (see also Boag 2005, 2012, 2017a, 2017b; Mackay and Petocz 2011b; Maze 1952, 1987/2009, 1993; Newbery 2011; Petocz 1999: 220ff.). This work challenges the standard view of Freudian drive as somehow divorced from cognitive and other psychological processes, and the related view that Freud struggled to integrate two competing theories of motivation—a drive-discharge model on the one hand and a psychological model on the other.

Freud's innovation centres on the intimate connection between the biological and the psychological, and is given its clearest account in what has come to be regarded as the *locus classicus* of his model of motivation—the famous passage in the *Interpretation of Dreams* (Freud 1900: 565–6) describing the hungry baby. This passage is variously cited as evidence for Freud's drive-discharge view, involving 'quasi-neurological or metapsychological details and speculations' (Pataki 2014: 33); as evidence for Freud's cognitive/psychological view, 'The basis for a cognitive, *as opposed to an instinctual* model of motivation' (Dahl 1983: 41, my emphasis); or as evidence for *both*, that is, for Freud's proposing 'side by side two quite different basic conceptions of motivation' (Dahl 1983: 40). But if we consider not only the details within the passage itself but also the material that immediately precedes and follows what is usually quoted, then Freud is identifying psychological states (e.g. perceiving, believing, remembering) as crucial *evolved components* of drive-consummatory activity, which 'makes use of movement for purposes remembered in advance. But *all the complicated thought-activity which is spun out from the mnemic image to the moment at which the perceptual identity is established by the external world—all this activity of thought* merely constitutes a roundabout path to wish-fulfilment which has been made necessary by experience' (Freud 1900: 566–7, my emphasis). Thus, the homeostatic needs of Freudian drives clearly *do* have access to the perceptual and cognitive machinery of consciousness, which has evolved to allow the drives to 'find' their objects in the external world; there is no disjunct between a *cognitive* and an *instinctual* motivational theory. Freud reinforces this position in his description of instincts and their vicissitudes (Freud 1915c), where he refers to an instinct's *appearing to us as* 'a concept on the frontier between the mental and the somatic, as the psychical representative of the stimuli originating within the organism and reaching the mind' (1915c: 121–2), regarding which Strachey offers the editorial suggestion that 'the contradiction is more apparent than real' (1915c: 113).[16] Moreover, Freud sees the (p. 267) evolved cognitive and perceptual apparatus as sustaining an elaborated drawn-out process of consummatory behaviour. He contrasts the reflex arc response to a mere stimulus with the far more complicated and developed response to instincts, in that they 'make far higher demands on the nervous system and *cause it to undertake involved and interconnected activities by which the external world is so changed as to afford satisfaction to the internal source of stimulation*' (1915c: 118–20, my emphasis). These 'involved and interconnected' activities include the various attachment and social behaviours. Freud's description here, when combined with his insistence on the

deterministic definition of instinctual drives, in which the 'desire' component of the standard desire-plus-belief model of action is transformed from the teleological *desire for* into the deterministic *driven by*, is reflected in Maze's paraphrase: the motivational structure's innately provided consummatory response is caused to run off and 'extended spatiotemporally to a tremendous degree by the fact of cognition' (Maze 1952: 38).

Moreover, Freud's commitment to the 'bodily ego' (Wollheim 1982) shows that it is not merely cognition that is accommodated in this more sophisticated drive theory but also cognition's agent, the ego or self, the executive that is widely supposed (e.g. in ego psychology, cognitive science, and cognitive motivation theory) to stand above the drives. Maze (1987/2009) identifies two strains in Freud's thinking—both appearing strikingly within a single paragraph of *The Ego and the Id* (Freud 1923b: 25): the ego-instincts view where the ego is 'that part of the id which has been modified by the direct influence of the external world', suggesting that the ego, being comprised of modified id drives, is instinctually based; and the ego as just a set of control functions, in which perception 'plays the part which in the id falls to instinct'. Yet, notes Maze, perception alone, furnishing policy-neutral factual information, cannot initiate action without that information being perceived as relevant by some existing motivation, and so both drive and perception acting together are necessary for the production of any action. Moreover, the self or psyche, being *rooted in* instinctual drives, is not 'passive'. As Maze, following Freud's (1923b: 23) reference to Groddeck, puts it, we 'are *lived by* our instincts' (1983: 176). In sum, disconnecting the ego or self from instinctual drives is not consistent with Freud's overall treatment of the ego, nor with other material in which he includes a Darwinian evolutionary acknowledgement of attachment theory:

> We have no reason to dispute the existence and importance of original, innate, distinguishing characteristics of the ego … we shall not overlook the fact that the id and ego are originally one; nor does it imply any mystical overvaluation of heredity if we think it credible that, even before the ego has come into existence, the lines of development, trends and reactions which it will later exhibit are already laid down for it.

(Freud 1937: 240)

(p. 268) Thus, Freud's motivational model sees the primary drives as foundational perceptually and cognitively based structures, together providing the mental plurality required to explain internal conflict.[17] They have evolved in the service of basic biological needs and in that same service have become extended into innate psychological social and attachment systems. The result is that everything we do, no matter how sophisticated and apparently removed from biological sources, is some form—either direct or elaborated—of drive-consummatory action (Gardner 1993; Maze 1983; Newbery 2011; Petocz 1999).[18]

With respect to the 'affect' component of the matrix, the emotional systems are, like cognition itself, evolved drive-ancillary structures via which the drives respond, using what James (1890: 471) called 'the bodily sounding board', to events pertaining to their

gratification/frustration. Just as there are no such things as non-cognition-connected drives, so there are no such things as non-affect-connected drives, and no such things as non-drive-based affects (e.g. if the hunger drive is frustrated, the organism may be thrown into an angry emotional state and behave accordingly). The only difference is that emotion is not an inevitable accompaniment or reaction. This is consistent with Wollheim's (1999) remarks that 'emotion is posterior to both belief and desire' (1999: 12) and that 'emotion rides into our lives on the back of desire' (1999: 15).

In the light of this, we can see contemporary affective neuroscience as contributing to fleshing out and implicitly confirming Freudian drive theory, albeit not without some muddying of the waters. Hopkins (2004) suggests that Panksepp's emotional command systems are 'roughly equivalent to what Freud meant by drives' (2004: 237). It seems to me that these systems confirm the motivation-cognition-affect matrix of Freudian drive theory while remaining uncertain on the role of drive, the question of primary drives, and the relations between drive, cognition, and affect. With respect to these systems Panksepp (1998) appears to acknowledge implicitly that drives are primary and that emotions are drive-ancillary. He says: 'emotional abilities initially emerge from "instinctual" operating systems of the brain, which allow animals to begin gathering food, information, and other resources needed to sustain life' (Panksepp 1998: 25). The drive-ancillary nature of emotional reactions is reinforced in the later Panksepp and Biven (2012) book, which begins: 'All of us get angry at times, *especially when our interests* (p. 269) *are ignored or thwarted*' (2012: ix, my emphasis), and ends with reference to the 'vastly interacting neural circuits that work in and for living bodies and that respond to the challenges of the world by creating desired circumstances and avoiding those that are harmful … emotional feelings are the experienced affective manifestations of such interactions' (Panksepp and Biven 2012: 500). Yet, when Panksepp discusses the origins of the emotional command systems and how they differ from drives, he clearly has in mind a Hullian-style theory of drives as abstract (and redundant) intervening variables (1998: 168), which is one of those traditional conceptions of drives that, as noted earlier, are inadequate and inferior to Freud's. This may explain why Panksepp says 'The most primal *affective*-cognitive interaction in humans … is encapsulated in the phrases "I want" and "I don't want" ' (1998: 27, my emphasis), adding examples such as dislike of bitter foods, or desire for food. These examples suggest that reference to *drive*-cognitive or *motivational*-cognitive interaction would be more appropriate. Thus, Panksepp's list of emotional command systems is, from the Freudian drive theory perspective, a heterogeneous mix of: general aspects of *all* drives (SEEKING); specific drives (SEXUALITY, CARE); behaviours (PLAY); and emotions (FEAR, RAGE, GRIEF). As Newbery (2011) points out, finding evidence of neural underpinnings of various behaviours is not sufficient for distinguishing an independent drive from its drive-ancillary structures and pathways or from other drives; an evolutionary criterion for identifying the primary drives is needed, and that is in any case implict in much of contemporary affective neuroscience, as also in Freud.

Given the standard perceptions of Freudian drive theory, it is not surprising that, as is evident from my earlier brief survey, most theorists who discuss symbolism, the primary process, and dreaming call for either abandoning or revising the Freudian metapsychology. There is a perceived gulf between Freud's (drive-based) metapsychology and his clinical theory, leading theorists to ask whether we should abandon the instinctual 'physiological level' of the metapsychology of classical Freudian drive theory in favour of the cognitive 'psychological level' of clinical theory. But Freud's metapsychology—keeping faith with his theory of the drive as a motivation-cognition-affect matrix—is more comprehensive than is typically claimed. While the unifying role of this metapsychology has been promoted by some authors (e.g. Aragno 1997, 2010; Kitcher and Wilkes 1988), there is dispute about its characteristics, and little acknowledgement of the crucial role of Freud's broader drive theory (but see Boag 2017b).

Further Thoughts Towards Clarification and Synthesis

Some of the questions listed in the first part of this chapter have been addressed indirectly during the revisiting of Freud's material, but we can now show more clearly the implications and how they might apply in clinical examples.

(p. 270) One major question concerns the contrast between regressive, defensive disguise versus progressive, transparent, healthy accommodation to reality. Freud's material suggests an integration that argues against too forced a contrast. His drive theory, underpinning the FB theory of symbolism, presents the mind as a society of competing impulses involving multiple often-conflicting drives seeking gratification via elaborated versions of their consummatory activities on their objects, and, where unavailable, substitute activities and substitute objects. Compromise formations are ipso facto motivation-cognition-affect accommodations to reality that also will typically include defensive elements; evasion of and accommodation to reality operate simultaneously within the complex situation. For example, in Freud's eighteen-month-old grandson Ernst's famous 'fort-da' game (Freud 1920: 14–16), there is no awareness of what the ritual symbolizes, but there are clearly both defensive and accommodative aspects to the 'game'. Similarly, that 'the dream is a slice of emotive-sentience, an expressive product of a feeling/thinking brain' (Aragno 2010: 11) is not incompatible with the fact that some of the material in that expression requires 'deciphering' by virtue of its distortions ('disguise'). Since the dream is neither a deliberate communication nor subject to deliberate disguise, the motivated mistaken identity that underlies dream symbolism arises from the same set of motivational structures that are also responsible for the dream's being a snapshot of the mind's emotional work, albeit while 'reality testing' is typically and inadvertently 'off-line'. This is consistent with the neuropsychological complexity-reduction account, in which the fictive or counterfactual active inferences in dreaming are simultaneously hallucinatory attempts at wish-fulfilment and involve psychological work and emotional processing, thus challenging Schneider's (2010) view

that Freudian dream theory does not allow for 'psychological work (yielding motivational advance)' (2010: 524). And, as Freud reminded us, not all dreams have a latent content. They can range from repression-free hallucinated satisfaction of earlier conscious but frustrated desire, as in little Anna Freud's dream about eating the strawberries that she had been earlier forbidden to eat (Freud 1900: 130), or existing desire, such as the pressing need to empty a full bladder, all the way to the dream fragment of the psychotic patient in which 'two halves join together' (a symbolic representation of desired reunion with the temporarily absent analyst) (Roussillon 2015: 587). By contrast, daydreaming and fantasy are hallucinatory experiences in which reality testing is deliberately or knowingly suspended. This is the sense in which (day)dreaming as emotional work goes on during waking. This is not to deny the significant difference between positions at either end of the continuum—between, on the one hand, the psychotic who is locked into the Segalian 'symbolic equation' (Segal 1958) and so is unable to play his violin in public and, on the other hand, the virtuoso who can exercise supreme control in sublimating sexuality and aggression into a public violin performance. It is, rather, to identify the common factors that are in play in both cases: motivational multiplicity and competition operating via the reality principle in the service of the pleasure principle with varying levels of conscious awareness and with evasion of reality and accommodation to reality operating simultaneously but applying to different aspects of the single complex situation. Symbolic play can involve the conscious and deliberate adoption of a substitute or (p. 271) 'pretend' object/situation with the suspension of disbelief that allows for unconscious treatment of the symbol as if it were the symbolized; it is a kind of motivated mistaken identity that involves a sliding scale of conscious awareness and unconscious absorption (as in Pataki's (2014, 2015) notion of *engrossment*). Adulthood sees the extension of symbolic play into the variety of healthy ego-syntonic accommodative adult forms (e.g. sport), forms of 'controlled regression', lucid dreaming, and so on. In sum, defensive conflict resolution is not incompatible with accommodation to loss and change, and in neither can *all* reality be avoided.[19]

Another major question concerns the relation of symbolism to psychosemiotics in general and to language in particular. To begin with, separating nonconventional symbolism from language allows us to see that all the references to interpretation, translation, communication, and so on pertain to the person who is attempting to understand ('interpret') what is going on—typically, to identify the symbolized. According to the FB theory, the presence of symbolism in dreams is nothing to do with language, so claims that there are two radically different theories of dreaming, as distinguished in the section on 'Dreams: contemporary developments and questions', conflate the dream's content with the tools used by the dream interpreter. Further, to present the 'mental digestion' of understanding and emotional processing as being an extension of narrow Freudian symbolism into the new 'symbolization' in which symbols are needed for basic thinking is misleading in two ways: it neglects prelinguistic mentality;[20] and it repackages what is already in Freud, that is, his clinical recognition of the emotional processing involved in successfully 'making the unconscious conscious'; since the sense of symptoms lies in some connection with the patient's experience, thinking that involves making links of

personal significance (Mackay 2003) is needed in order to 'make the unconscious conscious'; merely being informed by the analyst about the unconscious content produces no change 'until the conscious idea, after the resistances have been overcome, has entered into connection with the unconscious memory-trace' (Freud 1915a: 175–6).

We can see the combination of both of these points illustrated in clinical phenomena. For example, Heller (2013) refers to the famous wedding-night ritual of Freud's patient in his (1916/17) lecture on 'The sense of symptoms':

> This case is rich in its symbolic content, but also shows how the meaning of the ritual came to life once the woman could make the link between the obsessive ritual and what it stood for in her actual experience. Her *inability to symbolise* meant that the experience was repeated without any meaning for her. Once she linked it to her wedding night, she was able to create room in her mind for the symbolic meaning of her action.
>
> (Heller 2013: 238, my emphasis)

(p. 272) Here, 'inability to symbolise' does not mean that there is no symbolism of the FB motivated mistaken identity kind; she is of course symbolizing in just that way; instead, it is being used to mean 'inability to understand', via making the personally significant links, the actions that she feels compelled to execute repeatedly, and, a fortiori, inability to engage in the type of symbolization that is comprised of articulation of that situation.

Another example is seen in Taylor's (2010) discussion of the case of Parens's (2004) autobiographical *Renewal of Life: Healing from the Holocaust*. At the age of seventy-three, on the sixtieth anniversary of the murder of his mother at Auschwitz, Parens began to write the memories from his childhood, including his escape at age twelve from a camp in Vichy France. The writing was an excruciatingly painful process that led to the irruption of intense somatic reactions, including an eczema rash. Parens engaged in self-analysis, leading to his 'symbolization' of the symptom, in this case meaning that he came to discover its association with the lost memory of a coarse blanket. Notably, however, Parens is aware that this insight will not clear the rash away immediately because it needs to be emotionally *processed*: 'skin reactions of this nasty kind ... have to be, we might say, detoxified and metabolized by the psyche and the body' (in Taylor 2010: 188). Taylor says: 'As he was able gradually to tolerate and contain the dreaded affects, and link them with verbal representations as he continued thinking and writing, Parens presumably reduced the painful activation, which then allowed healing of his skin' (2010: 189). Here, Parens's ability to articulate and write about the 'meaning' of his symptom is dependent on his prior nonlinguistic ability to perceive the salient connection.

The last example comes in the famous autobiographical case study, written after her recovery, by a schizophrenic teenage girl (Sechehaye 1951), and with the analyst's technique of 'symbolic realization'. Hospitalized teenage 'Renee' had an obsession with (and accompanying psychotic fantasies about) apples, but would not eat those that were bought for her by her therapist/analyst Sechehaye (whom she called 'Mama'). Renee

reports: 'my eyes fell to her bosom, and when she insisted, "But why don't you want the apples I buy you?" I knew what I was yearning for so desperately, and I was able to bring out, "Because the apples you buy are food for grownups and I want real apples, Mama's apples, like those", and I pointed to Mama's breasts' (Sechehaye 1951: 70). Sechehaye understood, cut a piece from one of the apples she had bought and said 'Mama is going to feed her little Renee. It is time to drink the good milk from Mama's apples'. Renee continues: 'She put the piece in my mouth, and with my eyes closed, my head against her breast, I ate, or rather drank, my milk. A nameless felicity flowed into my heart' (1951: 70). Here, the analyst has provided the patient with the containment that allows her to suddenly identify the symbolic meaning of apples, and the analyst then engages in a safe and controlled way with that symbolic meaning, communicating it to her patient in a way that makes it emotionally digestible, by understanding the symbolism and explicitly engaging with it, in an action that is equivalent to symbolic play. Here, again, articulation and dialogue ride on the coat-tails of the FB theory's mix of nonconventional symbolism's motivated mistaken identity and its understanding and use by the therapist in symbolic play.

(p. 273) My final thought concerns the special nature of the Freudian contribution. The psychoanalytic field, like mainstream psychology, is characterized by what Rayner (1991) calls 'the empiricist suspicion of system-building' that 'can bring about neglect of coherent theory' (1991: 10). Yet that suspicion targets the type of system-building that was also rejected by Freud. His understanding of science and the kind of system that can be built via science was quite different. When addressing the question whether psychoanalysis offers its own *Weltanschauung*, he insisted that the scientific world view, of which psychoanalysis is a part, is to be contrasted with the kind of grand system offered by religion or Marxism, namely, 'an intellectual construction which solves all the problems of our existence on the basis of one overriding hypothesis' (Freud 1933: 158). That Freud could offer a genuinely integrative system from within the scientific perspective was made possible by his rejection of the dichotomy of *Natur-* and *Geisteswissenschaften*, and along with it dualism in its various guises (matter–spirit, free–determined, universal–particular) and various resulting pseudo-dichotomies (hermeneutic–causal, nomothetic–idiographic, individual–social, biological–psychological). While there is no doubt that the contemporary psychoanalytic scene involves a wide range of clinical practices, observational methods, general theories, and cross-pollinations with other disciplines, nevertheless, the view that we have 'moved on' from Freud closes the door prematurely. The insights and context-derived clarifications in his material can prevent unnecessary fragmentation. The special nature of his synthesizing contribution lies in the fact that it unifies and integrates without loss of due attention to complexity and diversity. In this, it leads the way for psychology more generally, and one of the main challenges for contemporary psychoanalysis is to show how that is the case.[21]

References

Aragno, A. (1997). *Symbolization: Proposing a Developmental Paradigm for a New Psychoanalytic Theory of Mind*. Madison, CT: International Universities Press.

Aragno, A. (2010). 'The bio-semiotic roots of metapsychology'. *Biosemiotics 3*(1): 57–77.

Bazan, A., Van Draege, K., De Kock, L., Brakel, L. A., Geerardyn, F., and Shevrin, H. (2013). 'Empirical evidence for Freud's theory of primary process mentation in acute psychosis'. *Psychoanalytic Psychology 30*(1): 57–74. doi:10.1037/a0027139

Bennett, M. R. and Hacker, P. M. S. (2003). *Philosophical Foundations of Neuroscience*. Oxford: Blackwell.

Bickhard, M. H. (1996). 'Troubles with computationalism'. In W. O'Donohue and R. F. Kitchener (eds), *The Philosophy of Psychology* (pp. 173–183). Thousand Oaks, CA: Sage. doi:10.4135/9781446279168.n13

(p. 274) Bion, W. R. (1962). *Learning from Experience*. London: Heinemann.

Boag, P. (2005). 'Addressing mental plurality: Justification, objections and logical requirements of strongly partitive accounts of mind'. *Theory & Psychology 15*: 747–767.

Boag, S. (2006). 'Freudian dream theory, dream bizarreness, and the disguise-censor controversy'. *Neuropsychoanalysis 8*(1): 5–16. doi:10.1080/15294145.2006.10773503

Boag, S. (2012). *Freudian Repression, the Unconscious and the Dynamics of Inhibition*. London: Karnac Books.

Boag, S. (2017a). 'On dreams and motivation: Comparison of Freud's and Hobson's views'. *Frontiers in Psychology 7*: 1–13. doi:10.3389/fpsyg.2016.02001

Boag, S. (2017b). *Metapsychology and the Foundations of Psychoanalysis: Attachment, Neuropsychoanalysis and Integration*. London: Routledge.

Brakel, L. A. W. (2004). 'The psychoanalytic assumption of the primary process: Extrapsychoanalytic evidence and findings'. *Journal of the American Psychoanalytic Association 52*(4): 1131–1161. doi:10.1177/00030651040520040201

Brakel, L. A. W., Kleinsorge, S., Snodgrass, M., and Shevrin, H. (2000). 'The primary process and the unconscious: Experimental evidence supporting two psychoanalytic presuppositions'. *The International Journal of Psychoanalysis 81*(3): 553–569.

Brakel, L. A. W. and Shevrin, H. (2005). 'Anxiety, attributional thinking, and the primary process'. *The International Journal of Psychoanalysis 86*(6): 1679–1693. doi: 10.1516/4CWJ-Q8R2-UDLP-HD5A

Brakel, L. A. W., Shevrin, H., and Villa, K. K. (2002). 'The priority of primary process categorizing: Experimental evidence supporting a psychoanalytic developmental hypothesis'. *Journal of the American Psychoanalytic Association 50*: 483–505. doi: 10.1177/00030651020500020701

Braun, A. (1999). 'Commentary on "The new neuropsychology of sleep: Implications for psychoanalysis" '. *Neuropsychoanalysis* 1: 196–201. doi: 10.1080/15294145.1999.10773260

Bucci, W. (1997a). *Psychoanalysis and Cognitive Science*. New York: Guilford Press.

Bucci, W. (1997b). 'Symptoms and symbols: A multiple code theory of somatization'. *Psychoanalytic Inquiry 17*: 151–172.

Bucci, W. (2008). 'The role of bodily experience in emotional organization. New perspectives on the multiple code theory'. In F. S. Anderson (ed.), *Bodies in Treatment* (pp. 51–76). New York: Analytic Press.

Capozzi, P. and De Masi, F. (2001). 'The meaning of dreams in the psychotic state: Theoretical considerations and clinical applications'. *The International Journal of Psychoanalysis 82*(5): 933–952. doi:10.1516/H4X5-52FM-8FWN-GQ52

Carhart-Harris, R. and Friston, K. (2010). 'The default mode, ego-functions and free energy: A neurobiological account of Freudian ideas'. *Brain 133*: 1265–1283. doi:10.1093/brain/awq010

Cassirer, E. (1923/1953). *The Philosophy of Symbolic Forms*, Vol. 1: *Language*, trans. R. Manheim. New Haven, CT: Yale University Press.

Cassirer, E. (1925/1955). *The Philosophy of Symbolic Forms*, Vol. 2: *Mythical Thought*, trans. R. Manheim. New Haven, CT: Yale University Press.

Cassirer, E. (1929/1957). *The Philosophy of Symbolic Forms*, Vol. 3: *The Phenomenology of Knowledge*, trans. R. Manheim. New Haven, CT: Yale University Press.

Christian, C. (2017). 'Review of: E. L. Auchincloss. (2015). *The psychoanalytic model of the mind*. Washington, DC: American Psychiatric Publishing'. *Journal of the American Psychoanalytic Association 65*(1): 127–133. doi:10.1177/0003065116684107

(p. 275) Cutler, S. E. and Brakel, L. A. W. (2014). 'The primary processes: A preliminary exploration of A-rational mentation from an evolutionary viewpoint'. *Psychoanalytic Inquiry 34*(8): 792–809. doi:10.1080/07351690.2014.968022

Dahl, H. (1983). 'On the definition and measurement of wishes'. In J. Masling (ed.), *Empirical Studies of Psychoanalytical Theories*, Vol. 1 (pp. 40–64). Hillsdale, NJ: Analytic Press.

da Rocha Barros, E. M. and da Rocha Barros, E. L. (2011). 'Reflections on the clinical implications of symbolism'. *The International Journal of Psychoanalysis 92*(4): 879–901. doi:10.1111/j.1745–8315.2011.00402.x

Deri, S. K. (1984). *Symbolization and Creativity*. New York: International Universities Press.

Dorpat, T. L. (2001). 'Primary process communication'. *Psychoanalytic Inquiry 21*(3): 448–463. doi:10.1080/07351692109348946

Fodor, J. A. (1975). *The Language of Thought*. Cambridge, MA: Harvard University Press.

Fodor, J. A. (1985). 'Fodor's guide to mental representation: The intelligent auntie's vade-mecum'. *Mind 94*(373): 76–100.

Fodor, J. A. (2008). *LOT 2: The Language of Thought Revisited*. Oxford: Oxford University Press.

Fotopoulou, A. (2012). 'Towards a psychodynamic neuroscience'. In A. Fotopoulou, D. Pfaff, and M. A. Conway (eds), *From the Couch to the Lab: Trends in Psychodynamic Neuroscience* (pp. 25–46). Oxford: Oxford University Press.

Fotopoulou, A., Pfaff, D., and Conway, M. A. (eds) (2012). *From the Couch to the Lab: Trends in Psychodynamic Neuroscience*. Oxford: Oxford University Press.

Freedman, N. (1998). 'Psychoanalysis and symbolization: Legacy or heresy?' In C. S. Ellman, S. Grand, M. Silvan, and S. J. Ellman (eds), *The Modern Freudians* (pp. 79–97). Northvale, NJ: Aronson.

Freud, S. (1895). *Project for a Scientific Psychology*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. I. London: Hogarth Press.

Freud, S. (1900). *The Interpretation of Dreams*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vols 4 and 5. London: Hogarth Press.

Freud, S. (1905). *Jokes and their Relation to the Unconscious*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 8. London: Hogarth Press.

Freud, S. (1910). 'The future prospects of psycho-analytic therapy'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 11 (pp. 141–151). London: Hogarth Press.

Freud, S. (1911). 'Formulations on the two principles of mental functioning'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 12 (pp. 218–226). London: Hogarth Press.

Freud, S. (1912). 'A note on the unconscious in psycho-analysis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 12 (pp. 260–266). London: Hogarth Press.

Freud, S. (1913). 'The claims of psycho-analysis to scientific interest'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 13 (pp. 165–190). London: Hogarth Press.

Freud, S. (1915a). 'The unconscious'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 166–215). London: Hogarth Press.

(p. 276) Freud, S. (1915b). 'Repression'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 146–155). London: Hogarth Press.

Freud, S. (1915c). 'Instincts and their vicissitudes'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 117–140). London: Hogarth Press.

Freud, S. (1916/1917). *Introductory Lectures on Psycho-analysis*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vols 15 and 16. London: Hogarth Press.

Freud, S. (1920). 'Beyond the pleasure principle'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 18 (pp. 7–64). London: Hogarth Press.

Freud, S. (1923a). 'Two encyclopaedia articles'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 18 (pp. 235–259). London: Hogarth Press.

Freud, S. (1923b). 'The ego and the id'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 19 (pp. 12–66). London: Hogarth Press.

Freud, S. (1933). *New Introductory Lectures on Psycho-analysis*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 22. London: Hogarth Press.

Freud, S. (1937). 'Analysis terminable and interminable'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 23 (pp. 216–253). London: Hogarth Press.

Freud, S. (1940). 'An outline of psycho-analysis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 23 (pp. 144–207). London: Hogarth Press.

Friston, K. (2010). 'The free energy principle: A unified brain theory?' *Nature Reviews Neuroscience 11*: 127–138. doi:10.1038/nrn2787

Gammelgaard, J. (2006). 'Primary process in metapsychology and cognitive psychology'. *Scandinavian Psychoanalytic Review 29*(2): 93–105. doi: 10.1080/01062301.2006.10592788

Gardner, S. (1993). *Irrationality and the Philosophy of Psychoanalysis*. Cambridge: Cambridge University Press.

Green, A. (1975). 'The analyst, symbolization and absence in the analytic setting (on changes in analytic practice and analytic experience)'. *The International Journal of Psychoanalysis 56*: 1–22.

Heil, J. (1981). 'Does cognitive psychology rest on a mistake?' *Mind 90*(359): 321–342, **http://www.jstor.org/stable/2253090**.

Heller, S. (2013). 'The capacity to symbolise: Ideas and implications'. *Psychodynamic Practice 19*(3): 236–250. doi:10.1080/14753634.2013.807979

Hibberd, F. J. (2005). *Unfolding Social Constructionism*. New York: Springer.

Hibberd, F. J. (2014). 'The metaphysical basis of a process psychology'. *Journal of Theoretical and Philosophical Psychology 34*(3): 161–186.

Hobson, J. A. (2009). 'REM sleep and dreaming: Towards a theory of protoconsciousness'. *Nature Reviews Neuroscience 10*: 803–813. doi:10.1038/nrn2716

Hobson, J. A. (2014). 'Introduction'. In N. Tranquillo (ed.), *Dream Consciousness. Allan Hobson's New Approach to the Brain and its Mind* (pp. 3–7). Cham, Switzerland: Springer.

(p. 277) Hobson, J. A., Hong, C., Carhart-Harris, R., and Friston, K. (2014). 'Virtual reality and consciousness inference in dreaming'. *Frontiers in Psychology 5*: 1133. doi:10.3389/fpsyg.2014.01133

Hobson, J. A. and McCarley, R. W. (1977). 'The brain as a dream-state generator: An activation-synthesis hypothesis of the dream process'. *American Journal of Psychiatry 134*: 1335–1348. doi:10.1176/ajp.134.12.1335

Holt, R. R. (1967). 'The development of primary process: A structural view'. In R. R. Holt (ed.), *Motives and Thought: Essays in Honor of David Rapaport* (pp. 345–383). New York: International Universities Press.

Holt, R. R. (2002). 'Quantitative research on the primary process: Method and findings'. *Journal of the American Psychoanalytic Association 50*(2): 457–482. doi: 10.1177/00030651020500021501

Holt, R. (2009). *Primary Process Thinking*. Plymouth: Jason Aronson.

Hopkins, J. (2004). 'Conscience and conflict: Darwin, Freud, and the origins of human aggression'. In D. Evans and P. Cruse (eds), *Emotion, Evolution, and Rationality* (pp. 225–228). Oxford: Oxford University Press.

Hopkins, J. (2012). 'Psychoanalysis, representation and neuroscience: The Freudian unconscious and the Bayesian brain'. In A. Fotopoulou, D. Pfaff, and M. A. Conway (eds), *From the Couch to the Lab: Trends in Psychodynamic Neuroscience* (pp. 230–265). Oxford: Oxford University Press.

Hopkins, J. (2015). 'The significance of consilience: Psychoanalysis, attachment, neuroscience and evolution'. In S. Boag, L. A. W. Brackel, and V. Talvitie (eds), *Psychoanalysis and Philosophy of Mind: Unconscious Mentality in the 21st Century* (pp. 47–136). London: Karnac.

Hopkins, J. (2016). 'Free energy and virtual reality in neuroscience and psychoanalysis: A complexity theory of dreaming and mental disorder'. *Frontiers in Psychology* 7: Article 922, 1–18. doi:10.3389/fpsyg.2016.00922

James, W. (1890/1950). *Principles of Psychology*, Vol. 1. New York: Dover Publications.

Jones, E. (1916). 'The theory of symbolism'. In *Papers on Psycho-analysis* (1948), 5th edn. London: Ballière, Tindall, and Cox.

Jones, E. (1953). *Sigmund Freud: Life and Works*, Vol. 1. London: Hogarth Press.

Kitcher, P. and Wilkes, K. V. (1988). 'What is Freud's metapsychology?', *Proceedings of the Aristotelian Society*. Supplementary Volume *62*: 101–115, 117–137.

Kövecses, Z. (2005). *Metaphor in Culture: Universality and Variation*. New York: Cambridge University Press.

Kris, E. (1952). *Psychoanalytic Explorations in Art*. New York: International Universities Press.

Langer, S. K. (1942). *Philosophy in a New Key*. Cambridge, MA: Harvard University Press.

Levy, R. (2012). 'From symbolizing to non-symbolizing within the scope of a link: From dreams to shouts of terror caused by an absent presence'. *The International Journal of Psychoanalysis 93*(4): 837–862. doi:10.1111/j.1745–8315.2012.00619.x

Mackay, N. (2003). 'Psychotherapy and the idea of meaning'. *Theory & Psychology 13*(3): 359–386.

Mackay, N. and Petocz, A. (2011a). 'Realism and the state of theory in psychology'. In N. Mackay and A. Petocz (eds), *Realism and Psychology: Collected Essays* (pp. 17–51). Leiden: Brill.

Mackay, N. and Petocz, A. (eds) (2011b). *Realism and Psychology: Collected Essays*. Leiden: Brill.

Malcolm, N. (1977). 'The myth of cognitive processes and structures'. In *Thought and Knowledge* (pp. 159–169). New York: Cornell University Press.

(p. 278) Maze, J. R. (1952/2009). 'Instincts and impulses'. In R. Henry (ed.), *Psychologies of Mind: The Collected Papers of John Maze* (pp. 25–40). London: Continuum Books.

Maze, J. R. (1983). *The Meaning of Behaviour*. London: George Allen & Unwin.

Maze, J. R. (1987/2009). 'The composition of the ego in a determinist psychology'. In R. Henry (ed.), *Psychologies of Mind: The Collected Papers of John Maze* (pp. 97–106). London: Continuum Books.

Maze, J. R. (1991). 'Representationism, realism and the redundancy of "mentalese" '. *Theory & Psychology 1*(2): 163–185. doi:10.1177/0959354391012002

Maze, J. R. (1993). 'The complementarity of object-relations and instinct theory'. *The International Journal of Psychoanalysis 74*: 459–470.

Maze, J. R. and Henry, R. M. (1996). 'Problems in the concept of repression and proposals for their resolution'. *The International Journal of Psychoanalysis 77*: 1085–1100.

McMullen, T. (2011). ' "Out there", not "in here": A realist account of concepts'. In N. Mackay and A. Petocz (eds), *Realism and Psychology: Collected Essays* (pp. 325–356). Lieden: Brill.

Michell, J. (1988). 'Maze's direct realism and the character of cognition'. *Australian Journal of Psychology 40*(3): 227–249. doi:10.1080/00049538808260045

Modell, A. H. (2014). 'The evolutionary significance of the primary process—The Freudian concept and its revision'. *Psychoanalytic Inquiry 34*(8): 810–816. doi: 10.1080/07351690.2014.968023

Newbery, G. (2011). 'Drive theory reconsidered (again!)'. In N. Mackay and A. Petocz (eds), *Realism and Psychology: Collected Essays* (pp. 839–871). Leiden: Brill.

Panksepp, J. (1998). *Affective Neuroscience: The Foundations of Human and Animal Emotions*. Oxford: Oxford University Press.

Panksepp, J. and Biven, L. (2012). *The Archaeology of Mind: Neuroevolutionary Origins of Human Emotions*. New York: W. W. Norton.

Pataki, T. (2014). *Wish-fulfillment in Philosophy and Psychoanalysis: The Tyranny of Desire*. London/New York: Routledge.

Pataki, T. (2015). 'Wish-fulfilment revisited'. In S. Boag, L. A. W. Brackel, and V. Talvitie (eds), *Psychoanalysis and Philosophy of Mind: Unconscious Mentality in the 21st Century* (pp. 1–46). London: Karnac.

Petocz, A. (1999). *Freud, Psychoanalysis and Symbolism*. Cambridge: Cambridge University Press.

Petocz, A. (2004). 'Science, meaning and the scientist-practitioner model of treatment'. In D. Jones (ed.), *Working with Dangerous People* (pp. 25–42). Oxford: Radcliffe Medical Press.

Petocz, A. (2006). 'Commentary on "Freudian dream theory, dream bizarreness, and the disguise-censor controversy" '. *Neuropsychoanalysis 8*(1): 48–53.

Petocz, A. (2011a). 'Re-thinking the place of semiotics in psychology and its implications for psychological research'. In S. C. Hamel (ed.), *Semiotics: Theory and Applications* (pp. 99–147). New York: Nova Science.

Petocz, A. (2011b). 'Why psychology has neglected symbolism and what a realist approach can offer'. In N. Mackay and A. Petocz (eds), *Realism and Psychology: Collected Essays* (pp. 597–651). Leiden: Brill.

Petocz, A. (2015). 'The scientific status of psychoanalysis revisited'. In S. Boag, L. A. W. Brackel, and V. Talvitie (eds), *Philosophy, Science and Psychoanalysis: A Critical Meeting* (pp. 145–192). London: Karnac.

Petocz, A. and Mackay, N. (2013). 'Unifying psychology through situational realism'. *Review of General Psychology 17*(2): 216–223. doi:10.1037/a0032937

(p. 279) Rapaport, D. (1951). 'Toward a theory of thinking'. In D. Rapaport (ed.), *Organization and Pathology of Thought* (pp. 689–730). New York: Columbia University Press.

Rayner, E. (1991). *The Independent Mind in British Psychoanalysis*. Northdale, NJ: Aronson.

Ridenour, J. M. (2016). 'Psychodynamic model and treatment of schizotypal personality disorder'. *Psychoanalytic Psychology 33*(1): 129–146. doi:10.1037/a0035531

Robbins, M. (2004). 'Another look at dreaming: Disentangling Freud's primary and secondary process theories'. *Journal of the American Psychoanalytic Association 52*(2): 355–384. doi:10.1177/00030651040520021201

Robbins, M. (2012). 'The primordial mind and the psychoses'. *Psychosis: Psychological, Social and Integrative Approaches 4*: 258–268. doi:10.1080/17522439.605905

Roussillon, R. (2015). 'An introduction to the work on primary symbolization'. *The International Journal of Psychoanalysis 96*(3): 583–594. doi:10.1111/1745-8315.12347

Schafer, R. (1976). *A New Language for Psychoanalysis*. New Haven, CT/London: Yale University Press.

Schneider, J. A. (2010). 'From Freud's dream-work to Bion's work of dreaming: The changing conception of dreaming in psychoanalytic theory'. *The International Journal of Psychoanalysis 91*(3): 521–540. doi:10.1111/j.1745–8315.2010.00263.x

Sechehaye, M. (1951). *Autobiography of a Schizophrenic Girl. With Analytic Interpretation*, trans. G. Rubin-Rabson. New York: Grune & Stratton.

Segal, H. (1958). 'Notes on symbol formation'. *The International Journal of Psychoanalysis 38*: 391–397.

Skinner, B. F. (1987). 'Whatever happened to psychology as the science of behavior?' *American Psychologist 42*(8): 780–786.

Solms, M. (2013). 'Freud's "primary process" versus Hobson's "protoconsciousness" '. *Contemporary Psychoanalysis 49*: 201–208. doi:10.1080/00107530.2013.10746545

Solms, M. and Turnbull, O. (2002). *The Brain and the Inner World: An Introduction to the Neuroscience of Subjective Experience*. New York: Other Press.

Solms, M. and Turnbull, O. H. (2011). 'What is neuropsychoanalysis?' *Neuropsychoanalysis 13*(2): 133–145.

Solms, M. and Zellner, M. R. (2012). 'Freudian drive theory today'. In A. Fotopoulou, D. Pfaff, and M. A. Conway (eds), *From the Couch to the Lab: Trends in Psychodynamic Neuroscience* (pp. 49–63). Oxford: Oxford University Press.

Tallis, R. (2011). *Aping Mankind. Neuromania, Darwinitis and the Misrepresentation of Humanity*. Durham: Acumen.

Taylor, G. J. (2010). 'Symbolism, symbolization, and trauma in psychosomatic theory'. In M. Aisenstein and E. Rappoport de Aisemberg (eds), *Psychosomatics Today. A Psychoanalytic Perspective* (pp. 181–199). London: Karnac Books.

Vanheule, S., Roelstraete, B., Geerardyn, F., Murphy, C., Bazan, A., and Brakel, L. A. W. (2011). 'Construct validation and internal consistency of the geometric categorization task (GEOCAT) for measuring primary and secondary processes'. *Psychoanalytic Psychology 28*(2): 209–228. doi:10.1037/a0022392

Vinocur Fischbein, S. and Miramón, B. (2015). 'Theoretical trajectories: Dreams and dreaming from Freud to Bion'. *The International Journal of Psychoanalysis 96*(4): 967–992. doi:10.1111/1745-8315.12269

Watt, D. F. (2012). 'Theoretical challenges in the conceptualization of motivation in neuroscience: Implications for the bridging of psychoanalysis and neuroscience'. In (p. 280) A. Fotopoulou, D. Pfaff, and M. A. Conway (eds), *From the Couch to the Lab: Trends in Psychodynamic Neuroscience* (pp. 85–108). Oxford: Oxford University Press.

Westen, D. (1998). 'The scientific legacy of Sigmund Freud: Toward a psychodynamically informed psychological science'. *Psychological Bulletin 124*(3): 333–71.

Westen, D. (1999). 'The scientific status of unconscious processes: Is Freud really dead?' *Journal of the American Psychoanalytic Association 47*: 1061–1106.

Wollheim, R. (1982). 'The bodily ego'. In R. Wollheim and J. Hopkins (eds), *Philosophical Essays on Freud* (pp. 124–138). Cambridge: Cambridge University Press.

Wollheim, R. (1999). *On the Emotions*. New Haven, CT/London: Yale University Press.

Yigael, Y. (2005). ' "The primary process": The vicissitude of a concept'. *International Forum of Psychoanalysis 14*(2): 76–85. doi:10.1080/08037060510030188

Notes:

([1]) That is, with careful terminological revision (e.g. 'implicit processing', 'affective schemas') and with no mention of Freud or psychoanalysis, thereby occluding important theoretical connections, such as those between Freud, Darwin, and the supposedly new attachment theory (see Hopkins 2004).

([2]) There are strong grounds (which I cannot go into here) for arguing that psychology itself is misguided in how it pursues these directions (see e.g. Bickhard 1996; Heil 1981; Hibberd 2005, 2014; Malcolm 1977; Maze 1991; McMullen 2011; Michell 1988; Skinner 1987; Bennett and Hacker 2003; Mackay and Petocz 2011a, 2011b; Petocz and Mackay 2013). And, since Freud led the way in *not* neglecting mentality, body, emotion, motivation, and evolution, there is no *conceptual* (as opposed to sociopolitical) need for psychoanalysis to detour through these mainstream 'corrective' movements.

([3]) This echoes the Cassirer–Langer post-Kantian view of the symbol as form of thinking (Cassirer 1923/1953, 1925/1955, 1929/1957; Langer 1942), but it also converges with the cognitive science acceptance of Fodor's (1975, 1985, 2008) 'language of thought' thesis, a thesis that, according to Heil (1981), 'rests on a mistake' (321).

([4]) Consistent with my earlier remarks, Vanheule et al. (2011) claim that the primary and secondary processes are both clinically useful and important for aligning psychoanalytic concepts with cognitive science and the neurosciences, and Ridenour (2016) commends Robbins's (2012) updating of Freud's primary process into the notion of the 'primordial mind' on the grounds that it integrates modern neuroscience findings and attachment research with Freud's basic formulation.

([5]) Hence, when Bucci later extends her cognitive approach to absorb the focus on emotion and motivation from second-generation cognitive science, she talks of 'emotional schemas' (Bucci 2008: 57) rather than drives. So also in Westen's (1998, 1999) championing of the psychoanalytic approach by assimilating it to the dominant connectionism in cognitive psychology: 'Was Freud wrong in some of his fundamental

ideas about human nature? Without doubt. His theory of drives was highly problematic' (1998: 362).

([6]) Hopkins's references to parallels, isomorphism, and consilience, and his glossing of 'integrate' as 'maps onto' is consistent with his caveat that the specific content of any particular theory (including Freudian) is not needed. Nevertheless, the point is that there is nothing in the neuroscience that disconfirms psychoanalysis.

([7]) This may explain Yigael's (2005) charge that Freud failed to explain 'why evolution would create such a "blind" information-processing system that is not as "smart" as the secondary process' (2005: 79).

([8]) The prominence of this dispute in the literature (see Boag 2017a for a recent discussion) confirms my earlier remark about the psychoanalytic embracing of contemporary neuroscience.

([9]) [Eds: For alternative, non-epistemic ways of distinguishing qualitatively between conscious and unconscious see the contributions to this Handbook by Fuchs, Phillips, and Finkelstein.]

([10]) According to this view, a naked thought must be *symbolized* in order for it to *be* a thought. This, as I mentioned earlier, became enshrined as the received view in cognitive information-processing theory of mind, where thinking is a symbol-using process. It is still the dominant view today, despite several cogent arguments against it, the most simple being that learning the first object-word-user relation requires prior perceiving/cognizing of those three terms (see references in footnote 2).

([11]) This is a clarification (prompted by Pataki, pers. comm.; see also Pataki 2015: 24–6) of my earlier somewhat misleading claim that 'it is an outcome of a mistaken belief in the identity of symbol and symbolized, a belief whose juxtaposition with a true belief in the *non*-identity of symbol and symbolized allows (via a process which is not properly understood) for the satisfaction of the repression-driven search for substitutes via compromise formations' (Petocz 1999: 195).

([12]) The issue of mental plurality is controversial, but it seems to me unavoidable given the facts of conflict and contradictory beliefs. One way of conceiving this is to link it to motivational plurality (see 'Freudian Drive Theory: A Motivation-Cognition-Affect Matrix' and footnote 15; also Boag 2005).

([13]) There is a big difference between identifying the third term of the triadic symbolic relation as a *cognizing organism* (as I do) and identifying it as an *interpreter*, because the latter implies linguistic or semiotic functions. There need be no interpreter or interpretation or communication; the organism simply acts towards X as if it were Y. It is for this reason that Freudian symbolism is not part of the Peircean theory of signs.

([14]) [Ed: For different appraisals of Freudian drive theory see contributions from Backstrom, Sandford, and Brook and Young, this volume.]

(^{15}) I am following standard psychological practice here in using 'affect' as an umbrella term covering emotion, feelings, and moods—as in Panksepp's (1998) 'affective' neuroscience that identifies 'emotional' command systems. Nevertheless, there is a vast literature involving considerable debate about these terms.

(^{16}) Freud claims that in mental life we know the instincts only by their aim; we are not perceptually connected to the biochemical processes of hunger (though we may feel the gnawing of the empty stomach), but we certainly immediately and urgently feel that we 'want food'. Desire, as Wollheim (1999) comments, is 'from the beginning of life, object-directed' (1999: 14; see also Maze 1993). Freud clearly did *not* share the view, espoused by many in cognitive neuroscience, and claimed by some to be an error in Freud's own drive theory, that consciousness is unrelated to homeostatic needs.

(^{17}) I earlier referred to the drives as the 'knowers', in that they have the 'ability to cognise, and thus be the subject term in psychological processes' (Petocz 1999: 222). But this lends itself to misunderstanding and to charges of postulating homunculi and falling into the mereological fallacy (Bennett and Hacker 2003); for it is, surely, the organism that perceives, knows, believes, etc. True. But when the organism is engaged in self-knowing, or when it is motivationally conflicted and epistemologically split—when, as in symbolism, the person simultaneously believes of the single object 'Here is X' and 'Here is Y', when in repression the person both wishes that p and remains unaware of the wish, when one part of the mind believes something while another part does not, then fine-tuning the single knower towards some kind of mental plurality appears unavoidable. In that fine-tuning, we can say that, while there is only one body, and only one set of perceptual apparatus, there are (*ex hypothesi*) several drives, and, accordingly, several organism-cum-particular-drive states, differentiated within the single bodily contour and the single shared perceptual apparatus.

(^{18}) In this sense the wide reach of the Freudian theory of symbolic substitutive activity via compromise formations overlaps considerably with the terrain identified by Pataki (2014, 2015) as Freudian wish-fulfilment (FWT).

(^{19}) The schizophrenic who refuses to play his violin in public is accurately perceiving the social sanctions against public masturbation, and so is accommodating to social reality. Moreover, he is acting rationally *given his false premise* about the nature of the violin.

(^{20}) See footnote 10.

(^{21}) I am grateful to Joel Michell, Marko Milic, Glenn Newbery, Tamas Pataki, and members of the Sydney Theory Group for helpful discussions on aspects of this material. My sincere thanks also to my research assistants, Ryan McMullan and Catherine O'Gorman. Finally, thanks to the editors, Richard Gipps and Michael Lacewing, for their thorough editorial work, including valuable insights and suggestions for improvement.

Agnes Petocz

School of Social Sciences and Psychology, Western Sydney University, Australia

Oxford Handbooks Online

Abstract and Keywords

The concept of wish-fulfilment as a substitutive mode of satisfying wish or desire was one of Freud's most important and singular psychoanalytic innovations. In his view dreams, neurotic symptoms, conscious and unconscious phantasies, delusions, hallucinations, jokes, much art, parapraxes, omnipotent thinking, illusions such as religion, the institutions of morality and social organization are all wish-fulfilling phenomena or attempts at wish-fulfilment. Although its remit is more circumscribed in contemporary psychoanalysis, wish-fulfilment can be seen to underlie such important conceptions as omnipotent phantasy, projective identification, actualization, and so on. This chapter exposes in detail Freud's singular innovation, relates it to some recent neuroscientific work, and shows how Freud's initial model of hallucinatory wish-fulfilment in dream and infant phantasy gratification, together with his conception of symbolism, can be extended to explain a range of symptoms as intentional behaviour, in line with Freud's ambitious claims for wish-fulfilment's remit.

Keywords: Freud, wish-fulfilment, symbolism, intentional action, self-solicitude

Ordinary Wish-Fulfilment and Freudian Wish-Fulfilment

When the seventeenth-century missionary Father Rageneau visited North America he found the Huron people distinguishing three causes of disease: natural causes; sorcery; unfulfilled wishes. Of the latter, some were known to the individual; others were not but could be revealed to him in his dreams. These pathogenic wishes could be concealed and then diviners were required to discover them. If the patient was mortally ill, the diviners would declare the objects of his wishes were impossible to obtain. If there were chances of recovery, a 'festival of dreams' was organized, a collection of valuables made among the tribe and presented to the patient amidst banqueting and public rejoicing. Patients not only recovered but sometimes emerged wealthy men.[1] Here is a perspicacious—almost 'Freudian'—diagnosis of the malady of frustrated (or ineluctable) desire, and a wonderful therapy. Here, pathogenic wishes, detected in dreams, are fulfilled in reality. It is important to see at the outset precisely how the ordinary conception of wishes being fulfilled instanced in this story differs from the topic of this chapter, Freudian wish-fulfilment (henceforth FW). In the ordinary conception of wish-fulfilment a person wishes for something; they get it; they know or believe that they get it and, as a result, cease wishing.[2] For their wishing to cease it is necessary that they at least believe that they get it: our cycles of desire and action are governed by information, not, or not typically, by exhaustion. And it is because they get it that they believe that they get it. Finally, it doesn't matter on this conception whether a wished-for object is attained by personal effort or by gift. Simplifying a little: where 'p' denotes any state of affairs:

> **(A)** In the *ordinary sense* of wish-fulfilment a wish that p is fulfilled only if:
> **(i)** the wish is terminated: the agent ceases to wish that p;
> **(ii)** the agent comes to believe that p;
> **(iii)** p is actualized;
> **(iv)** the wish is terminated *because* of the actualization of p.

The Freudian conception of wish-fulfilment is different. In the first edition of *The Interpretation of Dreams* (1900) Freud introduced a theory according to which dreams, neurotic symptoms, and conscious and unconscious phantasies, are wish-fulfilments. Eventually he concluded that positive psychotic symptoms—delusions, hallucinations—jokes and art, slips of the tongue, bungled actions, magical or omnipotent thinking, illusions such as religion, the institutions of morality and social organization were also wish-fulfilling or, recognizing that action can fall short of its object, attempts at wish-fulfilment.[3] The breadth of this array is striking: from the fictive representations of dream and phantasy to involuntary neurotic tics to the purposeful action sequences of religious observance, artistic creation, and character disorder. These phenomena, if indeed they are wish-fulfilling, are so in a manner patently different from the ordinary conception: in them, the objects of the wish are *not* attained, only a substitute or semblance is. FW is a

Subscriber: Tulane University; date: 16 December 2018

form of '*substitutive satisfaction*'. It is distinctive of the class that while conditions (i) and (ii) are necessary, conditions (iii) and (iv) are not. Characteristically in FW either the wished-for state of affairs is not actualized or, if it is, it has no causal role in the termination of the wish. The terminative role must therefore rest entirely with (ii) believing that p—or, perhaps, with experiential states functioning informationally and terminatively in the manner of belief. Moreover, with the possible exception of aesthetic appreciation, in all the cases described by Freud the agent initiates the wish-fulfilment. FW is something an agent does, either intentionally, subintentionally or by some expressive means.[4] It cannot be a gift. Thus, Freud's patient Dora sought by means of hysterical symptoms to separate her father from his mistress. Even if the symptoms were not manufactured intentionally by Dora, even if, say, they were dream-like expressions of desire or phantasy, they could still have succeeded in separating the lovers, and thus be wish-fulfilling in the Freudian sense. What would *not* count as FW is if, for example, a brick fell on the mistress's head.

On noting these features we may expect a spectrum of potential types of FW from simple non-intentional cases, to cases where intention plays some subsidiary role, to 'fully' intentional cases where FW is itself intended. To accommodate this spectrum we distinguish two conditions replacing conditions (iii) and (iv):

(B) For any wish that p, it is fulfilled in the manner of FW only if:

(i) the wish is terminated: the agent ceases to wish that *p* because:

(ii) the agent comes to believe (or, perhaps, experience) that *p*.

(v(a)) the agent initiates the wish-fulfilling process, in a sense that does not entail but does not exclude intention; or

(v(b)) in self-solicitous types, fulfilment of the wish can be truly described in some such way as 'A intentionally fulfils A's wish that *p*' or 'A intentionally gratifies (consoles, appeases) A'.

So, possible examples of the simple cases may be infantile hallucinatory wish-fulfilment and dreams, where representations caused by wishes in the primitive mental conditions Freud describes as primary process functioning achieve FW. Further along, representations of wished-for states of affairs may be generated intentionally, as in recalling fond memories or daydream, though not yet with express purpose of wish-fulfilment. In conditions of engrossment or reverie, however, vivid phantasy or recollection may ephemerally be mistaken for reality. At the furthest end of the spectrum are reflexive processes in which fictive representation, enactment, and manufactured or tendentiously selected 'evidence', is used intentionally for the purpose of fulfilling wishes. In the section on 'Self-solicitude' these reflexive, intentional forms of wish-fulfilment will be recognized as members of a class of self-solicitous activities, expressions of internal

object relations, that include self-consolation and care but also pernicious acts of self-appeasement and self-deception.

That is the conceptual outline of FW. It remains to show whether and how anything can coherently fall under the concept.

Freud's Basic Model

The concept of FW (though not of course articulated as stated earlier) is left undefined by Freud in several respects. He doesn't expressly distinguish FW from ordinary wish-fulfilment and the various conceptions of precursors such as the ancient oneiromancers[5]; and he often fails to distinguish between the *representation* of a wish being fulfilled, and the *fulfilment* of a wish (either as process or end state), though this distinction has significant consequences. Thus he writes of the famous Irma dream that its 'content was the fulfilment of a wish and its motive was a wish' (Freud 1900: 119) and of a 'delusion having as content the fulfilment of the wish' (Freud 1922: 226). But other passages go further: '*Dreams are … guardians of sleep which get rid of disturbances of sleep …* a dream does not simply give expression to a thought, but represents the wish-fulfilled as a hallucinatory experience … the dream does not simply reproduce the stimulus [the wish], but removes it, gets rid of it, deals with it, by means of a kind of experience' (Freud 1916–17: 129). Again: 'Symptoms serve for the patient's sexual satisfaction; they are a substitute for satisfaction of this kind …' (Freud 1916–17: 299). The idea that FW terminates or 'pacifies' (Hopkins 1995) wishes and provides substitutive satisfaction is in fact a theoretical necessity in Freud's scheme. Mere representation of wishes fulfilled fails to capture the specific functions assigned to dreams and symptoms: to preserve sleep, resolve conflict, circumvent realistic action. FW involves less than full satisfaction of a wish in the ordinary sense but more than mere representation: it involves, as well (usually temporary) termination of the wish.

Freud describes the simplest form of FW, the hallucinatory wish-fulfilment of the infant, in a famous passage:

> A hungry baby screams or kicks helplessly … the excitation arising from the internal need is not due to a force producing a *momentary* impact but to one which is in continuous operation. A change can only come about if in some way or other (in the case of the baby, through outside help) an 'experience of satisfaction' can be achieved which puts an end to the internal stimulus. An essential component of this experience of satisfaction is a particular perception (that of nourishment, in our example) the mnemic image of which remains associated thence forward with the memory trace of the excitation produced by the need. As a result of the link that has thus been established, next time this need arises a psychical impulse will at once emerge which will seek to re-cathect the mnemic image of the perception and to re-evoke the perception itself, that is to say, to re-

establish the situation of the original satisfaction. An impulse of this kind is what we call a wish; the re-appearance of the perception is the fulfilment of the wish; and the shortest path to the fulfilment of the wish is a path leading direct from the excitation produced by the need to a complete cathexis of the perception. Nothing prevents us from assuming that there was a primitive state of the psychical apparatus in which this path was actually traversed, that is, in which wishing ended in hallucinating. Thus, the aim of the first psychical activity was to produce a 'perceptual identity'—a repetition of the perception which was linked with the satisfaction of the need (1900: 565–6; cf. 1950: 317–19, 332–3).

First, there is the perception of a real experience of satisfaction achieved through 'outside help'. Subsequently when an internal need (or stimulus) gives rise to a wish (or psychical impulse) and their—presumably it is both the need and wish—immediate satisfaction is blocked, as it is in the helpless unattended infant and sleeper, a topographical regression, a backward movement of psychic energy to the perceptual system of the mental apparatus, ensues and a reactivation of a memory of satisfaction as a hallucinatory experience provides a 'perceptual identity' with 'an essential component' of the original experience of satisfaction. A memory of a previously satisfying experience is 'mistaken' for a perception of the original experience and it is for the infant *as if* its wish has been satisfied: the wish is terminated. This outcome is characteristic of the primary process mode of mental functioning whose key characteristics include mobility of cathexis and '*the replacement of external by psychical reality*' (Freud 1915b: 187). The hallucinatory or fictive experience of satisfaction (henceforth EOS) that terminates the wish does not terminate, but occludes, the endogenous stimuli that gives rise to the wish. Hunger gives rise to a wish to be fed; the infant hallucinates being fed and this pacifies the wish but not the hunger. Clearly pacification will be the more effective the more remote a wish is from the needs which give rise to it.

We are to imagine a similar process occurring in the rest of the wish-fulfilling series: dream, symptom, art, and so on. In essence, this is an informational conception of FW: wishing is terminated when a belief or, perhaps, an information-laden experience, registers that the wished-for state of affairs has been actualized. In the early *Project* (1950), Freud exposed an underlying energic or mechanical model on which the achievement of an EOS has several far-reaching neuronal consequences. One of them is a 'lasting discharge' which brings the 'urgency which had produced unpleasure in w [consciousness]' to an end, thereby reducing the psychical excitation of free (unbound) energy in accord with the overriding principle of inertia (or constancy) (Freud 1950: 318). The reduction (or elimination) of psychical excitation as the governing principle of mind remained a constant in Freud's metapsychological formulations. It is notable, however, even in this early work, that the energic conceptions ride on Intentional or common-sense-psychological ones; and this is of a piece with Freud's gradual realization that these idioms were indispensable for describing mind (cf. Freud 1895: 160–1). He realized, for example, that the regressive revival of a memory-percept in dream is by itself insufficient for the efficacy of the EOS; for the EOS to be efficacious the memory percept must, as well, 'meet with belief' (Freud 1950: 322–3, 339). Thus 'we could well imagine

the dream-work penetrating to mnemic images of this [regressive] kind, making conscious to us what was previously unconscious, and holding up to us a wishful phantasy which rouses our longing, but which we should not regard as a real fulfilment of the wish. Hallucination must therefore be something more than the regressive revival of mnemic images that are themselves Ucs' (Freud 1917: 231). So how do the mnemic images 'meet with' belief? In the state of sleep, Freud says, attention cathexis are withdrawn from 'the system Cs. as well as the other systems, the Pcs. and the Ucs., in so far as the cathexis in them obey the wish to sleep. With the system Cs. thus uncathected, the possibility of reality testing is abandoned; and the excitations which, independently of the state of sleep, have entered on the path of regression will find the path clear as far as the system Cs. where they will count as undisputed reality' (Freud 1917: 234). In the conditions Freud describes ideas or representations of wished-for states of affairs brook no contradiction and function as wish-fulfilling representations: psychical reality is mistaken for external reality. One might say that from the perspective of the dreamer the fictive EOS or hallucination functions as irrecusable evidence for wish-fulfilling beliefs. In such conditions, we may say, a person is *engrossed*.[6]

That is the basic model. We have noted that FW compasses as well as dreams and hallucination, neurotic and some psychotic symptoms, religious devotion, the making and experience of art, and so forth, but with the exception of 'symbolic wish fulfilment', understood in a wide sense, Freud provides little guidance on how the basic model can be extended to accommodate these other putatively wish-fulfilling phenomena. The latter generally involve intentional action, and it does appear to have been Freud's view, and I think it is the correct view, that some forms of FW are brought about intentionally, either as adventitious consequences of intentional acts or as directly intended i.e. where FW is itself intended in an act of self-solicitude. The basic model is evidently based on an incipient neuroscience that is theoretically meagre even for explaining dreams and hallucination, and especially so for the more complex cases of FW. The kernel of Freud's model can be extended, however, with considerable plausibility, as following sections show.

Representation and Symbolism

Apart from the transparent dreams of young children, few dreams, let alone symptoms and their kin, have conscious content that is manifestly wish-fulfilling. It is Freudian doctrine that these phenomena are mostly *disguised* wish-fulfilments, partly as an effect of unconscious primary processes and partly because of defensive activity to protect consciousness from disruptive stimuli. Freud distinguishes latent from manifest content in dream and symptom. In the case of dreams, the latent content or 'dream thoughts' consist of recent day residues interwoven with older aroused emotion-laden memories and wishes, and this content is transformed into the manifest, consciously experienced dream—the fictive experience of satisfaction or EOS (or cluster of EOSs)—by the

processes of the 'dreamwork': condensation, displacement, conditions of representability, secondary revision, and symbolism. So, famously, dreams require interpretation, and that ultimately involves uncovering the group of wish-fulfilments implicit in the dream: 'Dreams frequently seem to have more than one meaning. Not only … may they include several wishfulfilments one alongside the other; but a succession of meanings or wishfufilments may be superimposed on one another, the bottom one being the fulfilment of a wish dating from earliest childhood' (Freud 1900: 219). Of his 'specimen' dream of Irma's injection Freud writes:

> The dream fulfilled certain wishes which were started in me by the events of the previous evening (the news given me by Otto [that Freud's patient Irma was still unwell] and my writing of the case history). The conclusion of the dream, that is to say, was that I was not responsible for the persistence of Irma's pain, but that Otto was. Otto had in fact annoyed me by his remarks about Irma's incomplete cure, and the dream gave me my revenge by throwing the reproach back on to him … *I* was not to blame for Irma's pains, since she herself was to blame for them by refusing to accept my solution. *I* was not concerned with Irma's pains, since they were of an organic nature … Irma's pains had been caused by Otto giving her an incautious injection of an unsuitable drug—a thing I should never have done
>
> (Freud 1900: 118–19).

Freud's detailed analysis is too ramified to present here (see Hopkins 2015: 71–85; also Hopkins this volume) but the wish-fulfilments contained in the dream may be roughly divided into relatively obvious ones such as Irma's pains being organically caused, where the dream text practically wears the EOS on its sleeve ('Dr. M. … confirmed it … "There is no doubt it is an infection"'); and those that are disguised by the dreamwork. The 'incautious injection' with which Otto is rebuked leads Freud to partially repressed painful memories. Years earlier he had injected his patient Mathilde with a substance then considered harmless, resulting in her death; and on his advice Freud's friend Ernst began using cocaine for the relief of intractable pain, and by misuse quickly succumbed to it. So Otto with his thoughtless and probably dirty syringe is blamed not only for Irma's condition but also for the deaths of Mathilda and Ernst. If Otto is to blame then Freud is not. If FW is to succeed then it must be for the dreaming Freud *as if* Otto was responsible for the earlier deaths even though there is no conscious EOS to that effect. How does the dream do it?

At first pass, FW is effected by elements of the manifest dream representing multiple elements in the latent dream thoughts, utilizing condensation, displacement, and symbolism. Condensation leads chiefly to the formation of 'collective and composite images' of figures or actions. So we are told that Irma 'stood for' Freud's eldest daughter, his wife as well as Mathilda, and that she acquired a series of still 'other meanings' and 'allusions': 'Irma became the representative of all these figures which had been sacrificed to the work of condensation, since I passed over to *her*, point by point, everything that reminded me of *them*' (Freud 1900: 293). But how precisely does such representation by

means of condensation facilitate FW? Freud's unsatisfactory answer is that in condensation 'ideas are formed which are endowed with great intensity' and 'the intensities of the individual ideas become capable of discharge *en bloc* and pass over from one idea to another …' (Freud 1900: 595). This is not helpful, of course, because we have no idea how to relate *meaning* or representation to *psychical intensity*, or the facilitations effected by relations of meaning to discharge. Moreover, it turns on the awkward proposition that in condensation one element of a composite figure stands for other elements of it. Whatever we come to make of these problems we can see why Freud also considers displacement as a form of representation, for it too involves a transfer of 'psychical intensity' or 'value' (distinguished from 'sensory intensity'), in which an element of low value comes to stand for one with high value.

Condensation and displacement do not exhaust the relation of representation. Freud eventually recognized an exclusive type of representation—symbolism—attended with conditions distinguishing it from the other types. Symbols in this distinctively psychoanalytic sense are semantically opaque to the subject because their meaning or referent is either repressed or innately and unreflectively given, conditions not imposed on the other processes (Freud 1901a: 683). This is the sense in which a snake can symbolize a penis without condensation, or even displacement. Symbolism in this strict sense need not depend on these processes of the dreamwork.[7]

Putting aside Freud's unhelpful explanation in terms of energic discharge, can we explain how representation, either as symbolism (strictly) or in the awkward manner Freud attributes to condensation and displacement, facilitates and extends FW? Two considered theses recommend themselves. In regard to symbolism, Petocz (1999: 233) argued that the wish-fulfilling efficacy of the symbol rests on an unconscious belief in the identity of symbol and symbolized. As a consequence 'the symbol is mistaken for the symbolised and treated as if it were the symbolised' leading, 'for reasons not properly understood', to a kind of 'gratification which is not as complete as would be the gratification obtained from the satisfaction, via primary objects and activities, of the unopposed instinctual impulse' (Petocz 1999: 233). Such a belief in identity cannot be excluded (see the section on 'Intention in Wish-Fulfilment'), but it does not seem to be an indispensible condition for symbolic FW. It may be that an operation on the symbol is simply taken to be, is apprehended as, an operation on its referent, without belief in their identity.[8] The sorcerer who sticks pins in an effigy need not believe the effigy *is* his enemy. The supposed causal relations between effigy and referent are supported by an immense raft of beliefs and practices; there appear to be many different ways in which symbolic relations are established, and that may be true of unconscious symbolism also.

That reflection suggests the approach in Hopkins (2000; 2016) which bundles condensatory and (strictly) symbolic forms of representation under the head of symbolic cognition. It draws attention to the pervasiveness of symbolic and metaphoric associations informing all thought and action, naturally established over the course of life, so that 'present significant figures and situations partly inherit their meanings from past ones, to which they are unconsciously mapped' (Hopkins 2000: 12). So, given

Freud's fixed association between Irma and Mathilda, in representing Otto as giving a toxic injection to Irma, Freud at the same time pacifies his wish not to have been responsible for the death of his earlier patient. The pacification of the contemporary desire 'went together' with that of the deeper, more painful one (Hopkins 2000: 12). It would seem that for FW to succeed on either of these views—or on related views of unconscious representation—there must be *some* occurrence of unconscious understanding or apprehension underwriting them. Whether, as on Petocz's account, an unconscious belief in the identity of Irma and Mathilda entails that an operation on the one is experienced as if it were an operation on the other; or whether, as on Hopkins's account, representations of Irma and Mathilda become condensed and interchangeable through association; the pacification of Freud's wish that it was Otto (and not Freud) who had killed Mathilda would require that it be for him *as if* Otto had killed Mathilda. That would seem to entail *the acquisition of a belief that Otto killed Mathilda*—for there *is* no EOS whose content is Otto killing Mathilda. The passage from the EOS—the tableau which has Otto (almost) injecting Irma—to the content of Otto killing Mathilda requires a process of thought arriving at an unconscious understanding or belief that it is so.[9] So it would appear from the common-sense-psychological perspective. The symbolic associations between Irma and Mathilda may perhaps be best understood in sub-personal terms, but the wish-fulfilling belief about Otto killing Mathilda cannot be.

A Neuroscientific Model

The neuroscience upon which Freud drew was in its infancy and his conception of psychical energy was a placeholder waiting on the deliverance of more advanced science. In a series of papers Hopkins (2012, 2015, 2016, and this volume) has integrated fundamental aspects of Freudian theory with the recent 'free energy' neuroscience developed by Karl Friston and his colleagues, one of several neuroscientific models currently in contention. Friston's information-theoretic conception of free energy (FE) is not Freud's but, as Hopkins points out, they both assign FE the same overall functional role in the brain. In Freud's work (1950; 1911) and under Friston's (2010) FE principle the brain seeks to minimize FE aroused by sensory and interoceptive impingements. I summarize some relevant background to Hopkins's synthesis before focusing on the role the experience of satisfaction (EOS) plays in the account of wish-fulfilment that emerges from this neuroscientific model.

In the terms of this account (which bears similarity to Freud's notion of primary process and Kleinian unconscious phantasy), the brain comes equipped with an innate model of the world and a 'virtual reality generator' (Hobson, Hong, and Friston 2014), which generates predictions about sensory and interoceptive impingements and, by testing them in awake states, constructs increasingly veridical models of the world. A basic claim of FE neuroscience is that 'conscious perceptual experience is a form of explanatory hypothesis, that at once unifies, predicts, and inhibits the impingements that cause it,

thereby minimizing the FE associated with them' (Hopkins 2016: 3). The FE of a model is the difference between its complexity and its accuracy. *Accuracy* (roughly) is a measure of how well the model is predicting the data. *Complexity* reflects the number of parameters or potential trajectories for action composing the model and the extent to which they are modified in coming to predict data. Increase in complexity (without increase in accuracy) increases free energy. The imperative to minimize free energy therefore entails reduction in complexity, rendering the models more efficient in making predictions. During the day, it is argued, the huge influx of sensory input increases prediction error and, so, complexity. During sleep, with abatement in sensory input, and accuracy irrelevant, the brain or 'generative model' is freed to reduce complexity, enabling it to generate more efficient predictions in waking. It does this by, amongst other things, manufacturing in REM dream fictive or counterfactual virtual reality—various EOS—which inhibit realistic movement trajectories or parameters, including those associated with emotional and motivational conflict.

Hopkins makes the important suggestion that emotional and motivational conflict (too many or conflicting trajectories of action) and trauma (required emotional adjustments cannot be integrated) may plausibly be viewed on this model as sources generative of neurocomputational complexity. So in sleep, for example, the most significant nocturnal disruptions arise from recent emotion-laden memories of the sort Freud described as 'day residues' and those remotely linked and often disturbing emotions and wishes which are aroused and then condensed or consolidated with the more recent ones. Given the role of the brain as conceived above we can see how dreams and symptoms generic to phantasy or hallucinatory experience, which are viewed in psychoanalysis as ultimately products of conflict and trauma, may in large measure be productions functioning to reduce complexity. For if we now take the arousing materials of the dream as input, then, in accord with the basic claim of FE neuroscience, 'the fictive perceptual experiences of dreaming [i.e. experiences of satisfaction] could serve to unify, predict, and inhibit these arousals, thereby minimizing FE (as emotional conflict/complexity) in a way closely analogous to the realistic perceptual experiences of waking' (Hopkins 2016: 4). So on this view, the brain calculates the optimal motor trajectory for reducing free energy (on multiple interacting neurocomputational levels) by predicting *sensory trajectories* for courses of action. The EOS is then the culminating part of the sensory trajectory represented as conscious experience. But the EOS is not just the phenomenal end product of a sub-personal process; it is posited as having a causal role. In dream and symptom fictive experiences of satisfaction are figured as mitigating conflict, thereby reducing complexity and, so, FE, by eliminating incipient or conflicting trajectories or possibilities of action.

Precisely how does a fictive EOS reduce or eliminate conflicting parameters or potential trajectories? In essence this is a reformulation of the question of how the EOS is efficacious in FW. Recall the discussion of Freud's dream of Irma's injection. On Hopkins's view we have a sub-personal generative model acting in response to the dream

arousals 'inhibiting (pacifying) them by creating Freud's fictive experience of innocence and vindication, and so reducing the emotional complexity (adversity, conflict, and potential to cause depressive trauma) of the whole family of memories under reconsolidation' (Hopkins 2016: 14; this volume). The dream, he says:

> introduced an *ideal alternative counterfactual model* which perforce realized *an ideal fictive sensorimotor trajectory* by which, and in accord with Freud's priors of the previous day, the adverse emotional situation precipitated in him by Otto's unwelcome remark could be restored to a personally satisfactory (and free-energy minimal) equilibrium. Thus Freud could wake in the morning free from the self-justificatory and depressive trend in which he had spent the day and evening after Otto's remark ...' (2016: 15).

Now these passages hint at a number of perplexities that arise, I think, from the unresolved conjunction of sub-personal and personal perspectives. We can start with contrasting the view of dreams being exposed here with Freud's. According to Freud, the dream produces a transient EOS that pacifies disruptive wishes in a hallucinatory manner and guards sleep; it does not effect a lasting change in mind. It is perhaps in this vein that Hopkins wrote of the dream 'as a kind of *perfect and all-encompassing miniature hallucination*, in which the deluded dreaming subject utterly obliterates both what is happening in his mind and how things are in the world ...' (Hopkins 2012: 258). Here it is clearly engrossment in the EOS—the all-encompassing hallucination—that represses or at least pacifies parameters (conflicting trajectories, wishes etc.) inconsistent with the brain's dominant model. Wishes are pacified but they are not terminated. On the later account (Hopkins 2016; this volume), incompatible parameters are terminated or readjusted to achieve reduction in complexity. The dreamer receives not just a transient substitutive satisfaction enabling him to continue sleeping: his mind is permanently modified; thus, as Hopkins says, Freud awakes free from his depressive, self-accusatory mood to pursue the day with enhanced creativity. Perhaps that is so, or partly so. But then the generative model's work of emotional consolidation, conflict elimination, and inhibition as described by Hopkins at a sub-personal level seems to leave open the causal role, if any, of the Intentional and phenomenal properties of the EOS in complexity reduction. A situation where the generative model does indeed create an EOS, and the EOS does its work of complexity reduction, but entirely by virtue of its neural properties, not its Intentional and phenomenal properties, seems quite conceivable. In that case the *experience* of that complex tableau of EOS dreamt by Freud would seem to be epiphenomenal, causally inefficacious.[10]

If that is possible then there would appear to be ambiguity, if not incoherence, in this neuroscientific account of wish-fulfilment. But even if that is not right and there is no contradiction between what the sub-personal mechanism yields and the alleged Intentional causality of the EOS, the account would still seem to be deficient. As I argued in preceding sections, for a dream-wish to be fulfilled it must be for the dreamer *as if* the wish's object had been attained and *that* in most cases—given the absence of a

corresponding EOS—requires the *belief* that the object has been attained. It does not seem to me that the account of FW considered in this section satisfies that condition.

Whether or not that is the correct understanding of it, both the Freudian basic model and this more sophisticated neuroscientific one appear limited in accounting for the full range of FW as Freud envisaged it. Hopkins's thesis is directed at explaining dream, phantasy, neurotic and psychotic thought, and delusion. It is unclear (to this author, anyway) how it is to be elaborated to explain wish-fulfilling action, the extended ordered sequences of action constituting personality disorder, religious devotion, and aesthetic experience. The next two sections make an attempt to explain some of these phenomena along Freudian lines, retaining their presumptive wish-fulfilling character. But other, and perhaps more attractive, alternatives may present themselves: (i) to abandon Freudian wish-fulfilment as the fitting mode of explanation for these phenomena or (ii) to abandon altogether the attempt to provide Intentional explanation for the pathological developments Freud (and others) charted.[11]

Intention in Wish-Fulfilment

We said that the Freudian account of wish-fulfilment is intended to compass a wide array of activities including those that, on the face of it, are deliberate and intentional. The therapeutic imperative to formulate aetiologies for his clinical cases led Freud to hold: that neurotic symptoms, including enactments (obsessional routines, for example), are caused by unconscious wishes, reasons, and intentions; that these motives are not categorically different from corresponding conscious motives; and that they cause actions in much the same ways as conscious motives do.[12] That symptoms and their kin may have unconscious causes is relatively uncontroversial; the stronger contention that they may be produced intentionally, indeed as the end of practical reasoning, is problematic, and few have followed Freud so far. The view is counter-intuitive: symptoms certainly *appear* to be unintended, involuntary, and unpleasant. Again, *intention*, of all motivational concepts, has the closest affinity with conscious deliberation: 'unconscious intention' looks like an oxymoron. *Unconscious intentional agency* entails unconscious practical reasoning, acting, or resolving to act upon practical syllogisms—schemas whose main premise is a desire, minor premise an instrumental belief, and conclusion an intention or action. But as we shall soon see that is a taxing entailment. Moreover, given that in most typical behavioural symptoms conscious intentional agency remains intact, the operation of countervailing unconscious intentional agency would at least in some circumstances entail an apparently untenable partition or dissociation of self into quasi-independent, self-like centres of agency (Pataki 2003: 162–76).

Because of these difficulties most of the philosophers who reject partitive conceptions of mind as incoherent pursue non-intentional analyses of dream, wish-fulfilling phantasy and symptom; and largely ignore the role of FW in art, religion, and social action. My view, however, is that the concept of unconscious intentional agency is coherent and indeed essential for the explanation of the causal structure of many typical disorders and types of activity. To convincingly establish that view requires a larger canvas than I have here, but I can begin to sketch the main lines of its development.[13]

In Lecture 17 of the *Introductory Lectures* Freud discusses an obsessional symptom in a nineteen-year-old girl:

> [T]he most important stipulation related to the bed itself. The pillow at the top end of the bed must not touch the wooden back of the bedstead … The eiderdown had to be shaken before being laid on the bed so that its bottom end became very thick; afterwards, however she never failed to even out this accumulation of feathers by pressing them apart … She found out the central meaning of the bed ceremonial one day when she suddenly understood the meaning of the rule that the pillow must not touch the back of the bedstead. The pillow, she said, had always been a woman to her and the upright wooden back a man … If the pillow was a woman, then the shaking of the eiderdown till all the feathers were at the bottom and caused a swelling there had a sense as well. It meant making a woman pregnant; but she never failed to smooth away the pregnancy again, for she had for years been afraid that her parents' intercourse would result in another child and so present her with a competitor (1916–17: 265–8).

Freud understood this nightly ritual as the girl's wish-fulfilling attempt to prevent her parents' sexual intercourse and to resist anxieties connected with the birth of a competitor. The performance certainly presents as a sequence of intentional actions, but initially the girl could provide no reasons for performing them and found her own actions unintelligible. Before sleep anxious thoughts presumably assailed her: 'if I want to get some sleep then I had better separate the pillow from the bedstead etc.—I can't help doing that anyway'. But such thoughts do little to illuminate the causal structure of the symptom. Her performance is involuntary: against her conscious will. She doesn't know why she does what she does; she seems a mere spectator of her own actions. In not knowing her efficacious motives she doesn't consciously know *what* she is doing, under the descriptions that would be revelatory of why she behaves as she does. If we follow Freud's principal interpretation then as a preliminary attempt we can construct a practical syllogism for her behaviour:

> Girl wishes to separate mother and father.

> Girl believes that by separating bedstead and pillow she will separate mother and father.

> ---

> Girl separates bedstead and pillow.

The desire expressed in the major premise and the belief expressed in the minor premise are unconscious but if her performance accords with this syllogism it is plainly intentional. Then given her belief and action she will have achieved the Oedipal wish expressed in the major premise in the manner of FW: it will be for her *as if* she had separated her parents: her wish is pacified, anxiety is allayed, she sleeps. The glaring difficulty is that the minor premise expresses a 'mad belief': who could hold such irrational or bizarre beliefs, or act on them?[14]

We can in fact construct a number of practical syllogisms for the performance by considering again the issue of representation and symbolism. On the face of it the girl appears to unconsciously believe that if she separates the bedstead and the pillow she will have separated her parents. Believing this proposition is not the same as, nor does it presuppose, believing that her parents were identical with the furnishings (although the latter is not a belief that can be ruled out *tout court*). The sorcerer who sticks pins into an effigy doesn't believe that the effigy is his enemy. The distinction here is between one thing *symbolizing* another and that thing being *believed* to be identical with the other. And both of these can be distinguished from a third, where one thing is *equated* with another: where there is something like condensation or what Segal (1991) calls symbolic equation in which the indiscernibility between symbol and thing symbolized is such that an operation upon one is experienced as an operation on the other. So there are the following possibilities: if the girl believes unconsciously that the bedstead is her father and the pillow her mother then, from the fact that she separates bedstead from pillow she could unconsciously *infer* that she has separated her parents. In the case of symbolism, what the girl does effects—somehow—to separate her parents, as she apprehends it unconsciously. In the case of symbolic equation she again does not infer, she simply apprehends the performance as a separation of parents since operating on the furnishings is experienced by her as emotionally the same as operating on parents.

The practical syllogism set out above is an example using the sorcery-like belief 'if I separate the furnishings I will separate parents'. The following syllogism incorporates a symbolic equation:

> Girl wishes to prevent the conception of a baby-competitor.
>
> Girl believes that by keeping her father = bedstead and her mother = pillow apart she could prevent that conception.
>
> --
>
> Girl separates bedstead and pillow.

In each case (and also in the case of symbolism strictly), the girl performs an act rationalized and caused by her wish and the relevant belief, thereby intentionally fulfilling her Oedipal wishes in the manner of FW. It seems that each of these practical syllogisms articulates a possible instance of unconscious deliberation, and explains the ritual as an instance of FW. (Possible, that is, given certain psychoanalytic and philosophical premises undefended here.) But the case is underdescribed and we need not settle on a particular explanation. Moreover, there are still other possible schemas that do not trade on mad beliefs, symbolism, or symbolic equations, yet satisfy the requirements for FW as self-solicitude expressed in the opening section of this chapter. But before turning to those we are bound to examine the most cogent of the non-intentional alternatives.

It is objected that the rational reconstructions of the girl's behaviour considered here over-intellectualize a simpler situation. One important and influential approach was early advanced by Hopkins (1982, 1988). Hopkins placed FW at the centre of the common-sense-psychological explanation in the psychoanalytic domain but explicitly contrasted FW with intentional action. FW was conceived subintentionally, as supervening on imaginative activity animated by desire. 'In rational action', he wrote, 'motives produce willed intentions and real actions aimed at satisfaction. Here [in dreams] they produce wishes and mere representations of satisfaction, on the pattern of wishful imagining' (Hopkins 1988: 41). The subintentional mode of wishful imagining was also enlisted to explain symptoms such as the Ratman's obsessive thoughts of torture (Hopkins 1982: xxi–xxvii).

We may concede that, at the common-sense-psychological level, subintentional causation suffices for explaining the motivation of dreams, some phantasying, and symptoms such as the Ratman's obsessive thoughts. But it cannot fully explain more complex symptoms such as the Ratman's obsessive *actions*, the compulsive bedstead girl's rituals, or phenomena like wish-fulfilling religious practices. Suppose we try to explain the girl's performance as a product of wishing expressing itself *directly*, subintentionally, in action, as if it were dream, imagination, or phantasy? In that case no element of her action can be intentional: being caused subintentionally excludes a priori any action from having a reason or intention among its causal conditions. So if the performance *is* the phantasy, and the phantasy is the product of 'an exercise of will of a kind prior to that in intentional action' (Hopkins 1982: xxiv), then it cannot be intentional. But the performance has the hallmarks of being a sequence of intentional actions, features which this account is unable to explain.

A better interpretation of the proposal is this: the girl, agitated and *distrait*, separates bedstead and pillow and, given the (broadly, including condensatory) symbolic relations between parents and furnishings, she comes to *imagine* or phantasize unconsciously that she has separated her parents. Her actions are intentional, her wish-fulfilling imaginings are not (Hopkins 1982: xxv). On this attractive account the imagining and the consequent fulfilment of her wish supervene on an intentional performance that has acquired contingent symbolic meaning. Well, how exactly does the performance fulfil the girl's unconscious wish to separate her parents? On the view that a symptom is wish-fulfilling merely by virtue of being a representation of its originary wish-fulfilled the answer appears straightforward. But that view of FW has been rejected (in the section on 'Freud's Basic Model'): for representation alone, we saw, does not suffice for pacification and therefore for FW. The representation will pacify only if it is constitutive of the circumstance that *it is for her as if* her parents have been separated—a mental state that would seem to involve at least an unconscious belief or understanding that they had been separated. Without supplementation by some notion of unconscious understanding it is quite mysterious how the performance could acquire its pacific powers. *With* such supplementation we do arrive at FW: the girl does something: the act is imagined or understood unconsciously so as to generate a wish-fulfilling belief. But even so the account remains deficient. The wish-fulfilment is adventitious: the girl does something

which happens in the circumstances to pacify an unconscious wish. We do not yet have an understanding of *why she did what she did*. We seem driven to suppose that she did precisely what she did because she understood unconsciously that that very sequence of actions would achieve her aim. If not that, then two counter-alternatives suggest themselves: either she hit upon the wish-fulfilling performance by chance or her performance is determined fixedly by the symbolic associations established in the course of her life. Chance, I think we can exclude. On the other option we have it that given her fixed associations it is natural that she should so act; just as it is natural for the old bachelor to collect snuffboxes, given the unconscious meaning they have for him (cf. Hopkins 2000: 12). So, how do unconscious symbolic associations determine her performance? It is hard to say, but it seems that it cannot be *merely* through a fixed sub-personal association: for that would deprive the performance of its intentional character. On the other hand, if the symbolic associations do operate through the prism of unconscious understanding that would concede the present argument that she knew what the consequence of her performance could be.

However that uncertain issue is resolved, the subintentional approaches confront a further difficulty: the fact that pathological behaviour of very much greater complexity seems also to instance intentional FW. Some complex pathological enactments, for example wish-fulfilling character disorders, unfold over a lifetime. In such cases a series of otherwise unconnected acts, in diverse circumstances, strung out temporally, combine to achieve the one (say, self-defeating) purpose. A thousand acts pointing in the same direction. It seems inconceivable how an array of subintentional mechanisms, set of fixed symbolic relations, or chance actions could be strung together in the service of the one unified purpose: such behaviour seems to manifest resolution and plan. If such clinical cases cannot be explained non-intentionally then self-solicitous FW may offer some promise.

Self-Solicitude

In the first section of this chapter provision was made for cases of wish-fulfilment which fall under the heading of self-solicitude, cases in which the self takes itself as an object— mostly to be cared for, but allowing for the pathological permutations of such reflexive relationships that psychoanalysis has uncovered. We must now consider whether there are in fact psychological configurations that exemplify the self-solicitous forms of FW.

Granting a partitive conception of mind in which one part (B) is in solicitous relation to another, 'object' part (A), we can generate valid practical syllogisms for B:

B wants to console (protect, deceive etc.) A

> B believes that by separating bedstead and pillow A will be consoled (protected, deceived etc.)

> B separates bedstead and pillow.

Here B's belief is not mad; it is the kind of belief a concerned mother wanting to console or pacify a child may have. Construing the self as a unity of parts (in the manner indicated below) we recognize an expression of self-solicitude that was frequently described by Freud and other pioneer object-relations theorists.[15] Freud, for example, writes of the ego's protection and care-taking of the id (1940: 197–9) and of the superego's solicitude for the ego: 'To the ego … living means the same as being loved—being loved by the superego … The superego fulfils the same function of protecting and saving that was fulfilled in earlier days by the father …' (1923: 58); and 'if the superego does try to comfort the ego by humour and to protect it from suffering, this does not conflict with its derivation from the parental function' (1927: 166). That such self-solicitous relations are possible is a consequence, in Freud's view, of identification with objects resulting in structural differentiation shaping several features of the self[16]: the capacity to experience itself as divided; its capacity to divide, indeed for 'a real split' to occur between the agencies (1926: 97); and the capacity of parts of the self to then take other parts as objects. In this vein, Freud writes of the superego:

> The ego can take itself as object, can treat itself like other objects, can observe itself, criticize itself, and do Heaven knows what with itself. In this, one part of the ego is setting itself over against the rest. So the ego can be split; it splits itself during a number of its functions—temporarily at least. Its parts can come together again afterwards (1933: 58).

Freud's conception of the self is clearly partitive: the agencies are conceived as capable of entering into a range of mutual 'personal' relations, for example sadomasochistic relationships; they can function autonomously with their own perspectives, motives, and capacity for intentional action; they have a degree of mutual opacity. Despite these conditions Freud seems to have enjoined some kind of underlying unity in the self.[17] How might such extreme conditions of psychic stress be exemplified clinically? A patient:

> would rave against girl children and in fantasy would describe how she would crush a girl child if she had one, and would then fall to punching herself (which perpetuated the beatings her mother gave her). One day I said to her, 'You must be terrified being hit like that.' She stopped and stared and said, 'I'm not being hit. I'm the one that's doing the hitting'

> (Guntrip 1968: l9l).

This person certainly seems radically dissociated, a 'real split' having occurred (see footnote 15). There appear to be at least two internal perspectives: in the part doing the punching and in the part being punched. The parts appear to be mutually alienated: the speaker says she's not being punched, she's doing the punching. Guntrip observes that

his patient has identified with, and therefore adopted the perspective of, a punitive mother, while retaining the aspect of the detested child she at bottom feels herself to be, a 'child within' who maintains a masochistic relationship with an internalized mother. He mentions several motives: identifying with the punitive mother provides relationship with her; it represents the struggle to preserve an ego; it confers a sense of power, even if only over her self. Guntrip uses Fairbairn's conceptual framework which is more revealing of this clinical case, but Freudian language will do, and we can describe the girl's situation without excessive distortion as her ego being maltreated by a punitive superego that perceives its object as a wretched, hateful thing. On the other side, the ego perceives the superego under the aspect of the punitive, though still consoling, mother. These 'parts of the self' perceive and treat each other as if they were alien objects. The patient thus appears to be divided in perspective, subjectivity, and agency, in the terms of partition when a 'real split' has occurred.[18]

How is it possible for such an intrapsychic sadomasochistically configured enactment to be wish-fulfilling? Briefly, one way is this: in its enactment the superego figure is manufacturing an experience of satisfaction (EOS) for the ego—and at the ego's behest, or in response to its needs. We might say, paradoxically on the face of it, that the ego activates the superego and the superego then enacts the consoling/punitive repertoires written into it. The ego, not knowing fully the superego's repertoire, may get more than it bargains for. (Recall the quasi-independence and opacity of the agencies.) Principally the ego seeks the companionship and security provided by the internalized figure, and identification with it as aggressor, even if that comes with a familiar dose of punishment. In re-establishing that relationship, of which we can observe only the punitive enactment, the person, qua ego, has her wishes fulfilled, in the manner of FW. For we can see how the ego figure may be engrossed or immersed in the EOS that is the superego's enactment. So here are circumstances analogous to Freud's basic model of FW. There is, as it were, a theatre of irrecusable evidence presented to the ego from which it cannot but draw wish-fulfilling conclusions: 'mother is present and I am her beating a child'. The case of the bedstead girl may not be dissimilar. The ritualistic performance enacted as an EOS by an obsessional part of herself is unconsciously understood by her child-self as the wish-fulfilling achievement of separating her parents—while at the same time pacifying and achieving an unstable identification with that obsessional (superego) part.

W(h)ither Wish-Fulfilment?

Although Freud was educated in the reductionist biophysics movement of the late nineteeenth century even his earliest psychoanalytic work ran counter to the scientific positivism and neurological psychiatry associated with that movement. In contrast to it Freud insisted on the causality of subjective motives, on the desiring, wishful, and affective aspects of mind, particularly of course on sexual motives and their development. The idea that desire or unfulfilled wish causes mental conflict and disorder is older than

Plato and proved over the centuries, but with Freud something entirely new appears: mental disorder is not just *caused* by desires in conflict; symptoms (and their kin) are compromise *fulfilments* of those desires—in the manner of FW. This is a strikingly novel idea, fundamental to the explanation of human behaviour. Freudian wish-fulfilment, 'the doctrine of the wish' (as it was early on and enthusiastically designated (Watson 1916)), is a deep statement about the ineluctability of human desire and of the will to satisfy it. When we cannot change the world we change ourselves: if the world will not accommodate our desire then we dream, phantasize, manufacture delusions, modify our bodies and perceptions as bogus evidence that desire *has* been satisfied.

Desire and this human way of dealing with it are in some neglect in contemporary human sciences. Psychiatry, neuroscience, and cognitive psychology have attenuated their conceptual resources, so that it is unclear whether questions about mental illness once thought to be intricated with human longing and despair, or willed delusion, can even be framed in them. They are now dominated by forms of reductive materialism and simplistic conceptions of motivation. And were it possible to provide *complete* descriptions of mental disorder, dream, phantasy (let alone artistic creation and religious devotion) in the languages of the brain sciences then their exclusion of almost all traces of the Intentional, common-sense-psychological, understanding of disorder, of which FW has been an important component, would be entirely vindicated. But, of course, it is not.

In classical psychoanalytic theory most mental disorders *were* considered largely as pathologies of desire and conflict, described in the Intentional idiom; but that has given way in contemporary theory to a dominant concern with pathologies of developmental deficits resulting from environmental, attachment, or interactional failure. Psychoanalysis now appears less concerned with the pathology of desire than of belief—or the various proposed non-Intentional but belief-like structures anteceding beliefs—internal working models (Bowlby), representations of interactions generalized (D. N. Stern), and so on. The attention to environmental deficiency and early impairment has, by and large, diminished interest in the desiderative, conflictual, and wish-fulfilling aspects of mind. The theoretical contrast is odd because trauma, privation, deprivation become mensurable mostly against a backdrop of need and desire; and the language of object or interpersonal relations is otiose without reference to mutual desiderative and affective relations. So of course in reality motivational notions do not disappear and FW reappears, as omnipotent phantasy, projective identification, manic defence, and so on. The scope Freud had given it has been circumscribed but wish-fulfilment is still shifting in the background, wearing different clothes, too much unremarked.[19]

References

Boag, S. (2011). 'Addressing mental plurality: justification, objections and logical requirements of strongly partitive accounts of mind'. In N. Mackay and A. Petocz (eds), *Realism and Psychology*. Boston, MA: Brill.

Bollas, C. (1987). *The Shadow of the Object*. New York: Columbia University Press.

Bowlby, J. (1998/1980). *Attachment and Loss*, Vol. 3: *Loss*. London: Pimlico.

Brakel, L. A. W. (2009). *Philosophy, Psychoanalysis, and the A-rational Mind*. Oxford: Oxford University Press.

Cavell, M. (1993). *The Psychoanalytic Mind*. Cambridge, MA: Harvard University Press.

Cavell, M. (2006). *Becoming a Subject*. Oxford: Clarendon Press.

Dodds, E. R. (1951). *The Greeks and the Irrational*. Berkeley, CA: University of California Press.

Eagle, M. (2011). *From Classical to Contemporary Psychoanalysis*. New York: Routledge.

Ellenberger, H. F. (1970). *The Discovery of the Unconscious*. New York: Basic Books.

Freud, S. (1894). 'The neuro-psychoses of defence'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 3 (pp. 43–61). London: Hogarth Press.

Freud, S. (1895). *Studies in Hysteria* (with Joseph Breuer). In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 2. London: Hogarth Press.

Freud, S. (1900). *The Interpretation of Dreams*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vols 4 and 5. London: Hogarth Press.

Freud, S. (1901a). *On Dreams*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 5. London: Hogarth Press.

Freud, S. (1901b). *The Psychopathology of Everyday Life*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 6. London: Hogarth Press.

Freud, S. (1908a). 'Hysterical phantasies and their relation to bisexuality'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 9 (pp. 155–166). London: Hogarth Press.

Freud, S. (1908b). 'Creative writers and daydreaming'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 9 (pp. 141–153). London: Hogarth Press.

Freud, S. (1909a). 'Some general remarks on hysterical attacks'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 9 (pp. 227–234). London: Hogarth Press.

Freud, S. (1909b). 'Notes upon a case of obsessional neurosis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 10 (pp. 153–249). London: Hogarth Press.

Freud, S. (1911). 'Formulations on the two principles of mental functioning'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 12 (pp. 213–226). London: Hogarth Press.

Freud, S. (1912). 'Contributions to a discussion on masturbation'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 12 (pp. 239–254). London: Hogarth Press.

Freud, S. (1913). 'The claims of psychoanalysis to scientific interest'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 13 (pp. 163–190). London: Hogarth Press.

Freud, S. (1915a). 'Instincts and their vicissitudes'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 109–140). London: Hogarth Press.

Freud, S. (1915b). 'The unconscious'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 159–204). London: Hogarth Press.

Freud, S. (1916–17). *Introductory Lectures on Psychoanalysis*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vols 15 and 16. London: Hogarth Press.

Freud, S. (1917). 'A metapsychological supplement to the theory of dreams'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 217–235). London: Hogarth Press.

Freud, S. (1922). 'Some neurotic mechanisms in jealousy, paranoia and homosexuality'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 18 (pp. 221–232). London: Hogarth Press.

Freud, S. (1923). *The Ego and the Id*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 19. London: Hogarth Press.

Freud, S. (1925). *An Autobiographical Study*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 20. London: Hogarth Press.

Freud, S. (1926). *Inhibitions, Symptoms and Anxiety*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 20. London: Hogarth Press.

Freud, S. (1927). 'On humour'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 21 (pp. 159–166). London: Hogarth Press.

Freud, S. (1933). *New Introductory Lectures on Psychoanalysis*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 22. London: Hogarth Press.

Freud, S. (1940). *An Outline of Psychoanalysis*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 23. London: Hogarth Press.

Freud, S. (1950). *Project for a Scientific Psychology*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 1. London: Hogarth Press.

Friston, K. (2010). 'The free-energy principle: a unified brain theory?'. *Nature Reviews Neuroscience 11*: 127–138. doi:10.1038/nrn2787.

Gardner, S. (1993). *Irrationality and the Philosophy of Psychoanalysis*. Cambridge: Cambridge University Press.

Guntrip, H. (1968). *Schizoid Phenomena, Object Relations and the Self*. London: Hogarth Press.

Hampshire, S. (1974). 'Disposition and memory'. In R. Wollheim (ed.), *Freud: A Collection of Critical Essays*. New York: Anchor Books.

Hobson, A., Hong, C. C.-H., and Friston, K. (2014). 'Virtual reality and consciousness inference in dreaming'. *Frontiers in Psychology 5*: 1133. doi:10.3389/fpsyg.2014.01133.

Hopkins, J. (1982). 'Introduction: philosophy and psychoanalysis'. In R. Wollheim and J. Hopkins (eds), *Philosophical Essays on Freud*. Cambridge: Cambridge University Press.

Hopkins, J. (1988). 'Epistemology and depth psychology'. In P. Clark and C. Wright (eds), *Mind, Psychoanalysis and Science*. Oxford: Blackwell.

Hopkins, J. (1995). 'Irrationality, interpretation and division'. In C. McDonald and G. McDonald (eds), *Philosophy of Psychology*. Oxford: Blackwell.

Hopkins, J. (2000). 'Psychoanalysis, metaphor and the concept of mind'. In M. Levine (ed.) *The Analytic Freud*. London/New York: Routledge.

Hopkins, J. (2012). 'Psychoanalysis, representation, and neuroscience: The Freudian unconscious and the Bayesian brain'. In D. Pfaff, A. Fotopoulu, and M. Conway (eds), *From the Couch to the Lab*. Oxford: Oxford University Press.

Hopkins, J. (2015). 'The significance of consilience: psychoanalysis, attachment, neuroscience, and evolution'. In S. Boag, L. Brakel, and V. Talvitie (eds), *Psychoanalysis and the Philosophy of Mind: Unconscious Mentality in the 21st Century*. London: Karnac.

Hopkins, J. (2016). 'Free energy and virtual reality in neuroscience and psychoanalysis: a complexity theory of dreaming and mental disorder'. *Frontiers in Psychology* 7: 922. doi: 10.3389/fpsyg.2016.00922.

James, W. (1950 [1890]). *The Principles of Psychology*. New York: Dover Books.

Lear, J. (1998). *Open Minded*. Cambridge, MA: Harvard University Press.

Lear, J. (2015). *Freud*, 2nd edn. New York: Routledge.

Maze, J. R. (1983). *The Meaning of Behaviour*. London: Allen and Unwin.

Mele, A. R. (2009). *Effective Intentions*. Oxford/New York: Oxford University Press.

Moore, M. (1984). *Law and Psychiatry*. Cambridge: Cambridge University Press.

Newberry, G. (2011). 'Drive theory reconsidered (again)'. In N. Mackay and A. Petocz (eds), *Realism and Psychology*. Leiden/Boston, MA: Brill.

Ogden, T. H. (1993 [1986]). *The Matrix of the Mind*. New Jersey: Jason Aronson.

O'Shaughnessy, B. (1980). *The Will: A Dual Aspect Theory*. Cambridge: Cambridge University Press.

Panksepp, J. and Biven, L. (2012). *The Archaeology of Mind*. New York: W. W. Norton.

Pataki, T. (2000). 'Freudian wishfulfilment and subintentional explanation'. In M. Levine (ed.), *The Analytic Freud*. London/New York: Routledge.

Pataki, T. (2003). 'Freud, object-relations, agency and the self'. In M. C. Chung and C. Feltham (eds), *Psychoanalytic Knowledge*. Houndsmill: Palgrave.

Pataki, T. (2014). *Wishfulfilment in Philosophy and Psychoanalysis: The Tyranny of Desire*. London: Routledge.

Pataki, T. (2015). 'Wish-fulfilment revisited'. In S. Boag, L. Brakel, and V. Talvitie (eds), *Psychoanalysis and the Philosophy of Mind: Psychoanalysis in the 21st Century*. London: Karnac.

Pears, D. (1984). *Motivated Irrationality*. Oxford: Clarendon Press.

Petocz, A. (1999). *Freud, Psychoanalysis and Symbolism*. Cambridge: Cambridge University Press.

Rorty, A. (1988). 'The deceptive self: liars, layers and lairs'. In B. P. McLaughlin and A. Rorty (eds), *Perspectives on Self-deception*. Berkeley, CA: University of California Press.

Rosenfeld, H. (1987). *Impasse and Interpretation*. London: Tavistock.

Sandler, J. (1976). 'Dreams, unconscious fantasies and "identity of perception"'. *International Review of Psychoanalysis 3*: 1.

Sandler, J. and Sandler, A. M. (1978). 'The development of object-relationships and affects'. *International Journal of Psychoanalysis 59*: 2.

Segal, H. (1991). *Dream, Phantasy and Art*. London: Tavistock Routledge.

Solms, M. (2015). *The Feeling Brain*. London: Karnac.

Solms, M. and Kaplan-Solms, K. (2000). *Clinical Studies in Neuro-Psychoanalysis*. London: Karnac.

Spitz, R. (1965). *The First Year of Life*. New York: International Universities Press.

Steiner, J. (1993). *Psychic Retreats: Pathological Organizations in Psychotic, Neurotic and Borderline Organizations*. London: Routledge.

Watson, J. B. (1916). 'The psychology of wishfulfilment'. *The Scientific Monthly* Nov. 1916.

Weiss, J. and Sampson, H. (1986). *The Psychoanalytic Process*. New York: International Universities Press.

Winnicott, D. W. (1965). *The Maturational Process and the Facilitating Environment*. London: Hogarth Press.

Wollheim, R (ed.) (1974). *Freud: A Collection of Critical Essays*. New York: Anchor Books.

Wollheim, R. (1979). 'Wish-fulfilment'. In R. Harrison (ed.), *Rational Action: Studies in Philosophy and Social Science*. Cambridge: Cambridge University Press.

Wollheim, R. (1984). *The Thread of Life*. Cambridge: Cambridge University Press.

Wollheim, R. (1991 [1971]). *Freud*. London: Fontana.

Wollheim, R. (1993). *The Mind and its Depths*. Harvard, MA: Harvard University Press.

Wollheim, R. and Hopkins, J. (eds) (1982). *Freud: A Collection of Critical Essays*, 2nd edn. New York: Anchor.

Notes:

(1.) Father Rageneau, *The Jesuit Relations and Allied Documents* (1897–9), quoted in Ellenberger (1970: 26).

(2.) There is a sense of wish-fulfilment independent of belief and satisfaction, as when a deceased person's wishes are fulfilled after his Will passes probate. We can leave this sense aside.

(3.) For a summary see Freud (1913: 186–7). The 'actual neuroses' were early exceptions to the wish-fulfilment formula and later Freud made accommodations for traumatic dreams and symptoms, but the general view held consistently despite major changes to other aspects of theory.

(4.) The common-sense-psychological (CP) classification of motives and acts used in this chapter is as follows. *Intentional action* is caused and rationalized by the conjunction of desire and instrumental belief; it may also involve executive states such as intending, deciding, and choosing; but the having of reasons is both causally and logically necessary. Various non-intentional modes of causal explanation are available to CP. O'Shaughnessy (1980) highlighted the notion of *subintentional action*. Such acts, including mental acts, are caused by desire without facilitation by instrumental beliefs. There are forms of *expression* that have archaeology but no teleology, as when one clenches fists in (out of) anguish, or laughs. Explanation along these lines belongs in CP but is neither intentional nor subintentional. As Richard Wollheim and Jim Hopkins first pointed out, much of Freud's work can be seen as an extension and deepening of this common-sense psychological framework. Useful, if not unproblematic, classification of such extensions may be found in Wollheim (1991: xixff; 1993).

(5.) The dreams interpreted in the temples of Asclepius in which a respected personage or god, 'reveals without symbolism what will or will not happen, or should or should not be done' (Dodds 1951: 107), or those like Alexander the Great's Satyr dream (Freud 1900: 99), have a wishful tincture but they are obviously prognostic and instructive not wish-fulfilments or *substitutive satisfactions* in the Freudian sense.

(6.) Freud's implicit notion that an uncontradicted idea is *ipso facto* believed has pedigree (W. James (1950: II, 288) cites Spinoza as the source) but it remains live in the neuroscience of our day: 'In the absence of sensory constraints, the vivid percepts, delusional beliefs and cognitive defects cease to be delusional or defective—because these attributes are only defined in relation to sensory evidence. However, in sleep, there is no sensory evidence and the only imperative is to adjudicate and select among unconstrained scenarios that can be entertained by the sleeping brain' (Hobson, Hong, and Friston 2014). See also Solms (2015: 138–9). For elaboration of the notion of engrossment, introduced in Wollheim (1979), see Pataki (2014: 35–42).

(7.) Petocz (1999; this volume) distinguishes between Freud's narrow theory of symbolism (FN) of fixed archaic relations and a broader theory (FB). I accept that distinction. My concern here is to distinguish both these sorts of symbolism from condensation and displacement, vicissitudes that Freud also conceived as representational.

(8.) Petocz (this volume) modifies her earlier view and no longer contends that a belief in the identity of symbol and symbolized is essential to the motivated mistaken identity that underlies the capacity to treat one thing as if it were another. I present the earlier view because it still seems to me a live option, one of several that may facilitate symbolic FW, as the discussion in the section 'Intention in Wish-Fulfilment' shows (see also Pataki (2014): 88–94). Although she does not explicitly address the issue of symbolic wish-fulfilment in this volume, the argument advanced for the compatibility of drive discharge theory with 'cognitive' or 'informational' approaches to drive activity suggests than an account of FW in terms of substitutional drive discharge, where the symbol is mistaken for the drive's primary object, and the drive is figured as a complex entity with cognitive capacities, will suffice. My argument for the necessity in FW of the occurrence of a suitable EOS or wish-fulfilling belief suggests that it will not. In any case, I cannot follow Petocz (1999: 220–32; this volume; Maze 1983; Newberry 2011) in the constructions placed on Freudian drive theory. Maze, and in earlier work Petocz, argued that drives are programmed with innately, though highly modifiable, consummatory activities, that they must have the capacity to seek out and perceive the conditions that terminate drive activity, that these cognitive states are components of drives: in short, drives are 'knowers'. In the present chapter the relevant psychological states (perceiving, believing, remembering) are described as 'crucial *evolved components* of drive-consummatory activity', not as components of the drive itself, but a somewhat 'thick' view of drives is retained and it leads Petocz to favourably consider the implausible proposition that the psyche or self is 'composed' (exhaustively constituted?) of instinctual drives. These views seem to me neither accurate renderings of Freud's account of drives—ambiguous as that is—nor sustainable. The fact that drives causally interact with cognitive states—or that perceptions and beliefs influence our desires and behaviour—does not entail either that cognitive states are components, mereological parts of drives, or that drives can sense, peer at, or know anything. It must be conceded however that the 'thick' view of drives has the appeal of potentially resolving some of the many conundrums of mental plurality to which Petocz draws attention. (See also Boag 2011). I propose a different approach to some of these conundrums in the sections 'Intention in Wish-Fulfilment' and 'Self-Solicitude'.

(9.) This is discussed further in the section 'Intention in Wish-Fulfilment'. For reasons such as these Sandler (1976; Sandler and Sandler 1978) postulated an 'understanding work' running parallel with but in opposite direction to the distortions of dream and symptom. Thus the manifest or surface expressions are 'decoded' in the Pcs. of the topographical theory. or unconscious ego of the structural theory, and the latent meaning is then 'understood'.

(10.) I am assuming (the contested view) that the explanation of the causal efficacy of conscious Intentional and phenomenal properties qua mental properties remains problematic for non-reductive physicalism, the metaphysical view implicit in the work

under discussion. Conscious experience 'fits' in this enlargement of the Helmholtz-Friston account, and is made to do work, but is not, as it seems to me, explained.

(11.) Those possibilities are explored in Pataki (2014, 2015).

(12.) I use 'motive' loosely to designate anything that can be a cause of action. 'All the categories', Freud wrote, 'which we employ to describe conscious mental acts, such as ideas, purposes, resolutions and so on, can be applied to them [unconscious mental acts]. Indeed, we are obliged to say of some of these latent states that the only respect in which they differ from conscious ones is precisely in the absence of consciousness.' (Freud 1915b: 168). Freud invokes unconscious motives to explain the wish-fulfilling phenomena of symptoms (e.g. 1916–17: ch. xix–xiii, esp. 298–9; 1909a: 231); parapraxes (1901b); phantasies (1908a: 159, 164; 1909a); play (1908b: 146; 1911: 39); and art (1908b); daydreams and masturbation phantasies (1912). His description of neurosis as consoling flight into illness is certainly suggestive of intentional strategy, and it could hardly have been plainer that this was his view of the secondary purpose in neurosis (1909a: 231–2).

(13.) On the counter-intuitiveness of intentional symptom formation see Eagle (2011: 70). The intuition that 'intention is the one concept that ought to be preserved free from any taint of the less-than-conscious' (Hampshire, 1974: 125) is widely shared but currently much disputed (Mele 2009). Unconscious agency receives experimental support in Weiss and Sampson (1986) and neuropsychoanalytic support in Kaplan-Solms and Solms (2000: esp. 108, 111). For problems with partitive conceptions of mind see Gardner (1993). Some notable non-intentional readings of FW within a broadly common-sense psychological framework: Wollheim (1984, 1991, 1993), Hopkins (1982, 1995), Moore (1984), Lear (1998, 2015), Gardner (1993), Cavell (1993, 2006), and Brakel (2009). The larger canvas is in Pataki (2014, 2015).

(14.) [Eds: Cf. Leite (this volume) for further exploration of this theme.]

(15.) Spitz (1965: 179) noted that internalization of maternal functions is essential for the infant's capacity for self-regulation. Winnicott (1965: 48): 'The infant develops means for doing without actual care. This is accomplished through the accumulation of memories of care, the projection of personal needs and the introjection of care details'. To my knowledge the implications for mental structure of such normal internalization of caring functions has not been much investigated, though see Bollas (1987); Kaplan-Solms and Solms (2000: 234). More familiar are the pathological dissociations associated with the internalization of frustrating or hated objects, or with false-self formation based on compliance or denial. See Winnicott (1965); Rosenfeld (1987); Steiner (1993); Bowlby (1980). It is possibly because the self-solicitous forms of FW are so frequently dominated by pathological developments that they often issue, paradoxically, in self-deceit and symptom.

(16.) I use 'self' and 'parts of self' promiscuously, counting on a measure of disambiguation by context.

(17.) Freud: 'We were justified, I think, in dividing the ego from the id, for there are certain considerations which necessitate the step. On the other hand the ego is identical with the id, and is merely a specially differentiated part of it. If we think of this part by itself in contradistinction to the whole, *or if a real split has occurred between the two*, the weakness of the ego becomes evident … The same is true of the relation between ego and superego. In many situations the two are merged; and as a rule we can only distinguish the one from the other when there is a tension or conflict between them'. (1926: 97; my italics). Amongst philosophers, to my knowledge only David Pears (1984), Rorty (1988), and Pataki (2003, 2014) have argued for a conception of mind whose parts satisfy conditions for independent agency. Within psychoanalysis, on the other hand, it is commonplace to hold that the mind is constituted by independent agencies. Freud, Fairbairn, M. Klein, Winnicott, Bion, Ogden, and S. Mitchell explicitly embrace partitive conceptions of mind. A roll call doesn't settle anything about the adequacy of partitive conceptualization but it can suggest that the clinical phenomena press strongly in its direction.

(18.) I am of course taking a highly Realistic line on the agencies. The defence with a more accommodating conceptual framework is set out in Pataki (2014: ch. 5). Casting the agencies involved as person-like has the happy consequence that the language of personal relations, the Intentional or common-sense-psychological idiom, can be enlisted in the description of intrapsychic relations.

(19.) This chapter has been greatly improved by the meticulous commentary of the editors of this volume, and by discussions with Jim Hopkins and Agnes Petocz. I am much indebted to them.

Tamas Pataki

School of Historical and Philosophical Studies, The University of Melbourne

Oxford Handbooks Online

Integrating Unconscious Belief 🔓

Adam Leite

The Oxford Handbook of Philosophy and Psychoanalysis
Edited by Richard G.T. Gipps and Michael Lacewing

Abstract and Keywords

The very idea of psychic integration presents puzzles in the case of unconscious belief, both for the analysand and for the theorist. In many cases, the unconsciously believed proposition is one that the analysand knows perfectly well to be false. What could it be to bring such a belief to consciousness? What could psychic integration come to in this sort of case? Put bluntly, the task facing the analysand is to consciously hold the belief even while placing it within a broader perspective in which it is recognized to be false. Implications are drawn concerning a number of large issues in epistemology and philosophy of mind: Moore's Paradox, the role of rationality in psychic unity and self-consciousness, the nature of the first-person standpoint in relation to one's own attitudes, transparency accounts of self-knowledge, and the role of endorsement in the constitution of the self.

Keywords: unconscious belief, psychic integration, Moore's Paradox, rationality, self-consciousness, first person, self-knowledge, transparency, self-constitution

Adam Leite

> 'we find that unconscious beliefs control our patients' lives and they are therefore important to discern'
>
> Ronald Britton, as reported by Paul Williams (1996: 81)

Introduction

THERE are as many forms of psychic integration as there are forms of psychic disorganization and fragmentation. It's one thing to bring together feelings that have been held apart, another to take back projections, and still another to be able to represent previously unrepresented mental states or to bring psychosomatic phenomena more fully into one's mental life. My focus in this chapter is on just one form of

integration: the integration of psychodynamically unconscious belief. This issue looms large in clinical practice (Britton 1998: 8). It also has rich and surprising philosophical implications.

A step towards psychic integration is a step towards a more stable and more inclusive psychic organization from a previous disorganized, fragmented, or incomplete state. Stable organizations are possible that are wildly dysfunctional and out of touch with reality (Steiner 1993). Insofar as increasing integration is understood as a move towards psychic health, it must take place in a way that involves increasingly apt response to reality. At the ideal, 'integration' is a success term. It requires both getting things right enough about oneself and the world and also bringing these matters into contact with the rest of one's psychic functioning. In cases at issue here it involves insight, experientially based self-understanding (Bell and Leite 2016), self-acceptance, and acceptance of reality. A key question, then, forms the topic of this paper: What would all of this look like when it comes to certain psychodynamically unconscious beliefs that are—as the analysand can correctly see—plainly false?

(p. 306) My discussion will assume both a psychodynamic conception of the mind and a clinical picture on which integrating psychodynamically unconscious materials can be a step towards psychic health. My goal is to show what can be learned philosophically from taking this picture seriously. The discussion will be organized into the six following sections: 'Unconscious Belief' introduces the key notion under discussion; 'Integrating Unconscious Belief: The Problem' highlights a fundamental puzzle concerning the integration of unconscious belief; and 'Moore's Paradox and the Integration of Unconscious Belief' offers an initial solution. 'Is there a Superior Alternative Description?' considers and rejects various ways of redescribing the phenomena so as to avoid the puzzle. 'What Integration might Look Like' aims to make such integration intelligible, and, finally, 'Philosophical Implications' highlights important implications for several large issues in philosophy: the role of rationality in psychic unity, the nature of the first-person standpoint on one's own attitudes, the role of transparency in self-knowledge, and the place of endorsement in the constitution of the self. The upshot is a distinctive orientation towards these issues, one that may be said to be 'anti-rationalist' from the perspective of a common but overly narrow conception of what human rationality comes to.

Unconscious Belief

I will take for granted that there are beliefs which are unconscious for dynamic reasons relating to deep wishes, fears, and defences. For example, an analysand might unconsciously believe that he is the most delightful and most beloved of the analyst's patients and that their relationship is special in a way that even the analyst's marriage can't match. Another might unconsciously believe that her analyst has no interest in her life and struggles. These unconscious beliefs play a hidden role in shaping the person's conscious thoughts, desires, emotional responses, and behaviour. Conscious reflection

and deliberation generally will not modify them or otherwise affect the role they play. For these reasons, among others, it can be important for the analysand's positive development that such beliefs become conscious, in a sense of that term (discussed later) that is closely connected with self-awareness and the ability to self-ascribe the belief from a first-person standpoint.

In describing these states as *beliefs*, we take it that they are content-bearing states of the person that play a functional and explanatory role similar enough to that of ordinary belief to warrant lumping the two together. In characterizing them as *unconscious*, however, we highlight an important difference.

To bring the difference into view, it's helpful to have the notion of the *subjective perspective* of a belief. When somebody believes some proposition p, p's being the case is part of the subjective perspective through which and against the background of which the person encounters the world. We might say that in *her* subjective world, p is the case. Occupying this subjective perspective will involve patterns of emotional response and expectation, dispositions towards reasoning and action, certain phenomenal dispositions, **(p. 307)** and a tendency to seek corroborative evidence and to experience the world in ways congruent with p's being true.[1]

In ordinary conscious belief, the person is capable of occupying the subjective perspective of the belief *as conscious subject*. The world will be presented to and in her conscious occurrent thought as including *that p is the case*, and this presentation will include a distinctive phenomenology, which we might term the force of reality. In addition, ordinary conscious belief on the part of mature human beings involves *self-consciousness*, in the sense that the person can simultaneously *be* the conscious subject of the belief and also self-ascribe the belief from that very position. She can, we might say, occupy the attitude as her own.

In qualifying a belief as *unconscious*, then, we are highlighting that though the person is in a state which plays at least a good part of the functional and explanatory role of belief, the person currently either lacks the ability to occupy the subjective perspective of the belief as conscious subject or else lacks the ability to self-ascribe the belief from that position.[2]

For example, imagine someone with the following complex pattern of responses. He finds himself feeling guilty when kindnesses or favours are done for him, as though he has taken something away from the other person. He feels that he is in danger of being caught taking something he doesn't deserve if he accepts positive offers from others, and he experiences his legitimate and unproblematic requests for help as manipulative manoeuvres robbing others of their riches. He fears that accolades that come his way are nothing but the gains of trickery and deceit. He feels a vague terror that if others knew the truth about him they would hate and cast him out, so that he must always hide the truth about himself. It might be that what unifies all of these reactions is this: he believes that he is nothing but a thief. But though he has all of these thoughts and feelings both in particular situations and as part of the ongoing background of his life, he still might not

consciously encounter the world as containing this truth: *I am a thief*. In such a case, he cannot consciously occupy the subjective perspective of the belief, and he cannot self-ascribe it from that position. His belief is functioning unconsciously.

(p. 308) In understanding him in this way we attribute to him a belief content that conflicts *in his own thinking* with thoughts to the effect that he has not stolen. This too is supported by the detailed texture and particular pattern of his responses. For instance, a range of very particular emotional reactions (such as specific forms of guilt connected in various ways with the idea of theft) are forced upon him, while others (notably such attitudes as pride and contentment) are systematically excluded or distorted. He likewise interprets particular incidents in ways that fit with the thought that he has stolen, while conflicting facts are explained away, reinterpreted, or ignored. When told that he did well at something, he will react in a variety of ways that express, maintain, or even reinforce feelings of guilt connected with thoughts about theft. He might disavow that *he* really did the thing (insinuating that he stole credit for it) or deny that he did it *well* (insinuating that he somehow stole in accomplishing what he did). He might find a way to treat the compliment itself as stolen, e.g. by interpreting himself as having somehow 'fished' for it. As a last resort he might respond by defending against feelings of guilt and then feel that he is engaged in a cover-up. What he cannot do is sit comfortably with the thought, 'I just did well at such-and-such, and this person has, quite reasonably, voluntarily complimented me for it'. What we see here is a standing orientation towards himself involving feelings of guilt and thoughts about being undeserving, taking things from others, and the like, all of which he routinely connects with ideas of theft.

We gain a measure of understanding if we attribute a mental state to him that dictates these particular forms of response, an attitude that distorts his experience of himself and the world in ways that fit—and indeed insist upon—the truth of this proposition. This parallels theories of psychological attribution that emphasize interpretive fit with the overall pattern exhibited in the person's thought, feeling, and behaviour (cf. Child 1996; Davidson 1970: 221–2; 1983; 1984; Freud 1917; Mölder 2010). The pay-off is a gain in intelligibility, or as Freud terms it, 'a gain in meaning' (Freud 1915: 167).

Such psychological attributions are neither indefeasible nor unrevisable. In the therapeutic setting the attribution is underwritten by the ongoing fine details of the analys and's associations, thoughts, feelings, and reactions in interaction with the analyst. It might be refined retrospectively in light of later developments and the patient's settled understanding. In other cases, the state itself might change and develop over time.[3]

Many unconscious beliefs are unconscious for motivational reasons: that is, psychodynamic factors involving wishes, fears, and defences. The attribution of a particular psychodynamic explanation in a particular clinical case is supported in part by the way in which things unfold as the analytic process proceeds—for instance, defensive processes and patterns of resistance, the way in which the belief comes into consciousness (p. 309) when a particular defence or sense of unbearable conflict is

gradually undone, and the context provided by the ongoing development of a broader understanding of the patient's mind.

There are various routes by which unconscious belief can be psychodynamically generated. In some cases a recognized truth may be repressed because it is unbearably painful. In this way, traumatic real-world experiences can give rise to unconscious beliefs.

In other cases, a false belief may be psychodynamically generated. For instance, a belief may arise as a wish-fulfilment or as a defence against some other unbearable or unacceptable experience or psychic content. Britton (1995) describes a psychotic patient who believed that if she did not see her mother, she would go blind. This was a 'defensive counterbelief' formed to head off fear about what might happen to her mother when absent. (The protection comes from the childish thought, 'She's not absent; rather, something is wrong with my eyes and so I cannot see her'.)

Defensive counterbeliefs—like wish-fulfilling beliefs—can be conscious or can be kept unconscious because they are painful, threatening, conflict with central parts of one's conception of oneself or loved ones, and the like. In other cases, the belief is unconscious not so much to ward it off as to maintain it: the belief functions unconsciously as a way of safeguarding it in the face of acknowledged contradictory facts which are themselves felt (perhaps unconsciously) to be threatening or unacceptable.[4] Sometimes both factors are in play.

Importantly, psychodynamically unconscious beliefs (such as unconscious defensive counterbeliefs) need not be part of any larger rational structure. These beliefs exert rationalizing pressure insofar as they lead the person both to interpret situations in ways that conform with them and to fabricate justifications for convictions arising from them. But such considerations are not what holds them in place. Rather, they are held in place by motivational forces, mechanisms of defence, and other psychodynamic factors.

Two other important psychoanalytic notions—unconscious phantasy and implicit procedural schemata—stand in complicated relations to the concept of unconscious belief.

First, unconscious phantasy.[5] Different psychoanalytic traditions diverge in their interpretation of this protean concept (Bohleber et al. 2015). Frequently likened to daydreaming, phantasy is widely understood as 'an imagined scenario or storylike narrative' that represents the subject's interactions with others in ways 'shaped by motivational states and defensive operations' (Auchincloss and Samberg 2012a: 85). On many views, unconscious phantasy can lead to or produce both conscious and unconscious beliefs. There is disagreement, however, about several matters relevant to unconscious phantasy's relation to belief, including its origins, its relation to sources of motivation, (p. 310) and the nature of its content (propositional, conceptual, imagistic, or affective/bodily sensation).

Many prominent conceptions sharply distinguish the two. In a canonical Kleinian formulation, Susan Isaacs writes:

> phantasies are the primary content of unconscious mental processes … this 'mental expression' of instinct *is* unconscious phantasy. Phantasy is (in the first instance) the mental corollary, the psychic representative, of instinct. There is no impulse, no instinctual urge or response which is not experienced as unconscious phantasy.

> (Isaacs 1948: 91)

For Isaacs, unconscious phantasy is the first, 'low-level', mental manifestation of a drive or motivational urge ('instinct', in her terminology, following Strachey's translation of Freud's '*Trieb*' (drive, urge)). It likewise constitutes the mental content of defensive processes countering these urges. It has a non-propositional, affectively laden representational content (Gardner 1993: sect. 6.5). It thus differs from belief in several respects, including the fact that (unlike wish-fulfilling belief) it is not a *product* of the urge, but rather its mental expression.

Britton, too, sharply distinguishes unconscious belief from unconscious phantasy, holding belief to be what bestows the status of reality upon the phantasy's content for the subject (1995: 19–20; 1998: 9).

Writers in other traditions bring the two closer together. Erreich, for instance, suggests that unconscious phantasy should be 'defined as a more or less unconscious belief statement' (2003: 569) that results from defensive processes of compromise formation operating upon three components: wish, veridical perception, and naive cognition. However, Erreich also characterizes unconscious phantasy as 'a representational structure for mental content with a motivational component' (2003: 569) and views it as 'a vehicle for the mental representation of … affects, wishes, defenses' (2003: 545). So understood, it is doubtful that an unconscious phantasy simply *is* a belief, despite Erreich's explicit definition; it might perhaps be understood as a complex structure involving motivational elements along with unconscious beliefs as constituents, components, or 'concomitants' (2003: 567), or it might be understood as some sort of motivationally and affectively laden hybrid.

Some prominent views do not take a clear stand on the issue. In an influential contribution to the American ego psychology tradition, Arlow (2008/1969) characterizes unconscious phantasy as 'unconscious daydreaming' that is an 'ever-present accompaniment of conscious experience' (2008/1969: 42). If we take the identification with daydreaming literally, then unconscious phantasy is not a form of belief, even if it is 'composed of elements with fixed verbal concepts' (2008/1969: 23). Arlow also writes, however, that 'what is consciously apperceived and experienced is the result of the interaction between the data of experience and unconscious fantasying' (2008/1969: 42), and he likens unconscious phantasy activity to a source of input analogous to that (p. 311) provided by the senses (2008/1969: 43). He additionally characterizes unconscious

Subscriber: UC - Los Angeles (UCLA); date: 17 March 2019

daydreaming as a mental activity that 'supplies the mental set in which the data of perception are organized, judged, and interpreted' (2008/1969: 43). These formulations leave it unclear whether and to what extent unconscious phantasies are either belief-like or involve a belief component.

A second explanatory construct can be more clearly distinguished from the concept of unconscious belief. Some psychoanalytic theorists have recently emphasized nonlinguistic procedural schemas, models, or routines for interpersonal relating—psychological structures which are stored in an implicit (non-declarative) memory system and explain patterns of behaviour, thought, and emotional response by generating particular dispositions, motivations, and mental contents (including beliefs) in particular contexts (for instance, Beebe and Lachmann 1994, Fonagy 1999, Lyons-Ruth 1999, Stern et al. 1998; for discussion, Eagle 2013). Such schemata may be thought to be underlying psychological realizers of some unconscious beliefs. Whether that is so depends on larger debates in the philosophy of mind.

In any case, however, the two concepts are distinct and are not extensionally equivalent. First, many implicit procedural schemata—such as govern distance–standing as well as more complex interpersonal behaviour—are not psychodynamically unconscious. Rather, they remain difficult to articulate because stored in implicit (non-declarative) procedural memory (Fonagy 1999: 216–17; Stern et al. 1998: 905–6). They consequently have different behavioural manifestations. The patterns of resistance, defensive response, and shifts over time that will lead an analyst to attribute a psychodynamically unconscious belief are present only in certain cases. Moreover, as noted earlier, some unconscious beliefs are generated by processes of psychic defence. Some of these can have a very different sort of clinical manifestation and profile, and can play a different explanatory role, from that of implicit procedural schemata. The two categories are thus disjoint. However, none of this is to deny that procedural schemata for interpersonal relating can participate in psychodynamic processes and structures: for instance, they can sometimes generate psychodynamically unconscious phenomena, remain unacknowledged for psychodynamic reasons, be shaped by psychodynamic processes, and even involve characteristic interpersonally enacted defences (Lacewing, this volume).

Integrating Unconscious Belief: The Problem

Much clinical theorizing emphasizes that it can be crucially important for a patient's progress that unconscious beliefs become conscious. As Britton puts it, 'until a patient discovers that he really believes an idea its correspondence, or lack of it, with external reality is irrelevant. His belief remains suspended, unconscious, unmodified, and unverified, in parallel with a contrary belief' (1995: 21).

(p. 312) What matters here is not simply awareness that one has a particular belief. An analysand might be said to 'discover that he really believes an idea' in the sense that he learns to recognize the signs of his unconscious belief as they show up in his thought, speech, and behaviour. On this basis he can say, 'I believe that so and so'. However, such

a person remains no closer to his belief *as conscious subject* than he was before (Moran 2001; Finkelstein, this volume). This might be a stage in a positive development nonetheless, perhaps even a step towards bringing the belief into consciousness. The person might benefit from learning to predict and manage the effects of the belief even if it never becomes fully conscious. Still, this is fundamentally different from being able to consciously occupy the subjective standpoint of the belief and self-attribute the belief from precisely that position. It is only when one has attained that latter ability that one can self-reflectively recognize the belief in operation 'from the inside' and bring it into contact with one's ongoing thinking about the matters with which it is concerned.

Bringing an unconscious belief into consciousness does not merely involve coming to occupy this position in relation to that single belief. Rather, it involves bringing the particular belief into the larger system of one's conscious thoughts, feelings, beliefs, and desires. It thus both requires and constitutes a process of psychic integration—a process of bringing together into a single perspective what has been held apart or cut off, for instance in a process of psychic defence. This necessarily involves an increase in the range of positions—thoughts, feelings, and beliefs—that one can simultaneously self-consciously occupy as a conscious subject.

To approach what might be involved here, it can be helpful to consider a parallel schematic example. In the defence known as 'splitting', conflicting emotional responses—along with correlative contradictory representations of self and other—are kept distinct (Auchincloss and Samberg 2012b; Kernberg 1966: 238). The defence operates to evade the anxiety introduced by the felt conflict between the emotions, and it can be exacerbated by the analysand's sense of threat arising from intense feelings of hatred (Klein 1946). If an analysand has split her feelings of love and hatred, alternately experiencing herself as in loving communion with an ideal object and in fierce battle with a denigrated object, it will be crucial for her positive development that she come to experience her hatred, understanding it as such, *while also remaining in contact with her loving feelings*. That is the work of integration in this sort of case. The difficult task facing the analysand is to hold in mind the feelings of both love and hate, understood precisely as love and hate, and to experience their conflict as such, thereby uniting them into a single self-conscious perspective. This amounts to bringing her loving self and her hating self together into one.[6]

Integration in this sense does not necessarily entail rational control. Someone who manages to integrate her love and hatred may not be able to get rid of either by reflecting on the reasons for and against them. She may not be able to modulate the one by 'mingling' (p. 313) it with the other (and that may not be appropriate in any case). Still, her new ability to hold both in mind will enable new forms of thought, feeling, and reaction, including the possibility of experiencing her behaviour and motivations in new and emotionally richer ways, asking new questions about herself and others, and considering paths of action that formerly were not live options. This requires the ability to

recognize the conflict without becoming overwhelmed by anxiety, so that she can live within and from that position of conflict as a self-conscious subject.

Turn now to the case of dynamically unconscious belief. In many cases the belief is one that the analysand would perfectly well recognize as *false*.[7] For instance, consider again the belief *I am a thief*. Despite the reactions manifesting this unconscious belief, the person might know perfectly well that he is no thief and has earned his salary and received loving gifts and assistance from family and friends. Here, concretely, is what this might look like. If we ask him, 'Have you earned your salary?', he replies, fully sincerely, 'Yes, of course. I work very hard.' And yet he might feel an inarticulate sense of guilt and then begin telling a story about someone else who embezzled. If he is given a gift out of love, he might recognize the gesture and genuinely feel grateful. And yet he might also find himself feeling that he has gotten away with something. These feelings and thoughts do not do away with his knowledge any more than everyday obsessive worries do away with one's knowledge that one turned off the stove. In his self-conscious thinking and reasoning, he retains a firm grip on the truth.[8]

Such cases present a particular challenge within the lived experience of the analysand. The analysand must attempt to take up, as a conscious belief, something he knows to be plainly false. There is a difficulty here which is not merely a matter of defence against anxiety or psychic pain. But these cases also present puzzles for us as theorists. What would it *be* to integrate a plainly false belief? How can we make sense of the idea of such a belief's entering into the person's self-conscious subjective perspective?

First, integration does not entail elimination. It would be lovely, perhaps, if as soon as the troublesome belief began to come into consciousness it automatically ceased to exist. However, because of the belief's roots in wish, phantasy, and defence, this is not how things usually go. These motivational forces require a different sort of treatment. For the same reason, integration does not guarantee modifiability through deliberation. It is a mistake to suppose that as the belief enters self-conscious mental life, it must become directly and fully accessible to control, modification, or revision through conscious reflection (contra Parrott 2015; cf. Moran 2001).

At the same time, however, integration can't require radical loss of one's grip on reality. It *can't* be that a belief comes into consciousness only insofar as the person now self-consciously believes, without qualification, something that she previously consciously (p. 314) regarded as false. Consider, again, Britton's patient who believed that if she did not see her mother, she would go blind (1995: 22). Such beliefs exert a rationalizing force of their own, leading to the construction of a structure of delusional reasons and evidence. To accept such a structure without qualification, and to feel and act accordingly, wouldn't be a route to integration, but to psychosis—a severe break from reality (as in the case of Britton's patient). Likewise, though confusion is sometimes an important aspect of the process of working through, it would not be a positive development in this sort of case if the person's ongoing conscious state were to become one of genuine uncertainty or confusion about what is true.

For these reasons, to understand what it would be to integrate such a belief we need to make room for an initially surprising idea: that a person could *simultaneously* self-consciously believe something *and* recognize that the belief is false. That is, we have to allow for the possibility of at least a period in which the person recognizes that she believes that she has stolen everything good that has come her way—and occupies the subjective position of that belief, viewing the world in that way—even while also perfectly well recognizing that this belief is false because there are a great many good things in her life that she did not steal: some were earned through hard work, some were freely given, and sometimes she was simply lucky. This state may not be the ultimate goal—an issue to which I will return—but we have to allow for it as a possible and sometimes essential step along the way.

Moore's Paradox and the Integration of Unconscious Belief

What I have been describing is a psychological position that might be summarized in an utterance or affirmative judgement of the so-called 'commissive' version of Moore's paradox: namely, 'I believe that p, but p is false' (Moore 1993). Many contemporary philosophers take such an utterance or affirmative judgement to be paradoxical. While it can be true that someone believes something that is false, it can seem incoherent to self-ascribe a belief from the first-person standpoint even while judging or declaring that the believed proposition is false. This incoherence has seemed to many philosophers to go beyond mere irrationality; it is, we're told, a fundamental and defining feature of the concept of belief. As Wittgenstein writes, 'One can mistrust one's own senses, but not one's own belief' (2001 [1953]: 190), and 'If there were a verb meaning "to believe falsely", it would not have any significant first person, present indicative' (2001 [1953]: 162). Many recent philosophers (e.g. Coliva 2015; Heal 1994; Moran 2001) have followed Wittgenstein in holding that we can make no sense of the idea of someone's consciously occupying the subjective standpoint of a belief, self-attributing it from that position, and also acknowledging its falsity. And yet this is precisely what integration of dynamically (p. 315) unconscious beliefs often seems to require. The psychoanalytic case thus requires us to accept as actual precisely what this tradition deems unintelligible.[9]

To make sense of this, it's helpful to think more about what is involved in belief.

Consider the difference that is made when a belief functions in the way characteristic of conscious belief. Conscious belief that p constitutively involves that p's being the case is part of the subjective perspective through which and against the background of which the person consciously encounters the world. As noted earlier, the world will be presented to and in her occurrent thought as including *that p is the case*, and this presentation will include a distinctive phenomenology: a sense or feeling of reality that might be captured by saying that it is part of her subjective world that things are this way.[10] Her experience will likewise be shaped in various ways that fit with the truth of p, and she will be disposed to have congruent emotional and motivational responses. Central here too are

other differences in the person's dispositions. If she consciously believes that p, she is disposed to judge that p when she considers relevant questions. P will be presented in her conscious directed thinking as a premise for her reasoning about what to do and what is the case, and she will be disposed to so deploy it. She will be disposed to assert p when appropriate and to act on the basis of p.

All of this can be true even if a person also recognizes that her belief is false. Crucially, many of these dispositions are defeasible, and they can be defeated in various ways and to varying extents in different cases. In particular, someone might consciously believe that p, seeing the world in just the ways involved in consciously having this belief, and yet not deploy p as a premise in conscious reasoning, act on its basis, or assert it flat out. She might be disposed to judge p to be true and yet at the moment when she considers the question, judge it to be false. She might strive not to act in ways that are predicated on taking p to be true. Some of this may involve resisting these dispositions through an act of will, as when one consciously resists the strong temptation to act as if p were true. In other cases it may simply be a matter of what the person does in the course of her directed conscious thinking: she might simply think, 'But p is false', and that might be enough. These dispositions can also be defeated through changes in the person's underlying psychological functioning, so that the temptations and tendencies arising from these dispositions do not even show up in her conscious thinking.

(p. 316) Suppose, then, that the analysand recognizes both that there is in fact no evidence in favour of the thought that she is a thief and that all sorts of considerations support the contention that she isn't one. Suppose that she also understands that her belief *I am a thief* is itself a product of dynamic processes relating to deep wishes, fears, and defences, and suppose that she has some experientially based understanding of these wishes and fears and her defensive responses to them, so that all of this is functioning to some extent consciously as well. In such a case, the belief will be intelligible to her, no matter how wacky it is, in a way that goes beyond merely intellectual understanding, insofar as the belief links with emotions and motivations she has experienced 'from the inside'. In light of all this, she might experience herself and her world in all of the ways involved in believing that she is a thief, and yet she might attempt not to reason from *I am a thief* in her conscious thinking about what to believe, think, and do. She might not assert 'I am a thief', full stop. She might attempt to recognize when her emotional responses are born of this false belief, and she might attempt to treat them accordingly; likewise, with motivations that are given life by the belief. All of these further responses, along with the subjective perspective involved in believing that p, thus form a total complex perspective in which she self-consciously believes that she is a thief while not endorsing that belief and indeed while judging it to be false. Even as she sees the world through the lens of this belief, as it were, she also recognizes it as her lens, recognizes the ways in which it is inapt, and resists, suppresses, or otherwise does not engage in some of the patterns of response that would otherwise appropriately be involved.

Different psychological positions might meet this description, as I discuss later. For now, the point is this: we can in fact make sense of someone aptly capturing the reality of her overall psychological situation by simultaneously self-attributing a belief from the subject position and rejecting that belief as false. The key is to see that even as one occupies the subjective perspective of the belief as conscious subject, one's *total self-conscious perspective* at that moment might include a wider purview. While this is a rationally incoherent position, situations of this sort are familiar in ordinary life as well, particularly in cases in which one unexpectedly discovers that a view to which one is deeply attached is in error (Leite 2016).[11]

Any view of belief which emphasizes its connection with characteristic patterns of activity, thought, motivation, and the like, would lead us to expect exactly the sort of phenomena that I am pointing to. At least, it would do so unless one also holds the view (identified most prominently with Donald Davidson) that belief ascription requires that the person come out as *rational*, since I am proposing belief ascriptions that involve (p. 317) contradictory beliefs within the same total view. However, if we think of belief ascription as part of the attempt to *make sense of the person* or to *render her intelligible* —where we separate that notion from specifically *rational* intelligibility—then often precisely such a belief ascription will be the best fit. The distinctive form of intelligibility provided by person-level folk-psychological explanation is not primarily and narrowly rational intelligibility (Hursthouse 1991); a man's rubbing his face in his wife's sweater is made intelligible when we note that he is dealing with her recent death, and this intelligibility is not fully captured when we attempt to understand this case in terms of a 'rationalizing explanation' in the standard narrow sense. It rather has to do with considerations of meaningfulness and significance for the person: with the emotional resonances things have, the ideas with which they are associated, the motivations that are in play, and the like. Once we allow that these richer forms of intelligibility play a role in the attribution of attitudes such as belief, the possibility opens up of attributing much more complex, conflicted belief states to the person.

To sum up: I have been arguing that the formerly dynamically unconscious belief is integrated by being incorporated into a total conscious perspective that also contains its contradictory. The total perspective is thus rationally fractured. But the fracture is contained, in three senses. First, it is *limited*, in that it does not metastasize into a larger split that would introduce a division into the person's capacity for thought (as in compartmentalization or dissociation). Second, the fracture is contained insofar as the person is able to *hold the whole situation in mind*—both sides of the fracture, the fracture itself, and their relations, and so can experience the whole situation as an object of thought and as less threatening. Third, the various aspects of the person's overall perspective—including the formerly unconscious belief—are put in their places: they are given a distinctive organization in relation to each other. As a result, the person now functions differently. The person's overall response—given the contradiction—is as rational as it can be, insofar as the person aims not to act or reason on the basis of the false belief and will attempt to ameliorate its effects on her emotions and motivations.

This is the total complex situation that might be summed up and expressed by the utterance or thought, 'I believe that p, but p is false'.

It should be emphasized that this is *not* the mechanism that philosophers sometimes call 'fragmentation' (Lewis 1982). In fragmentation, two (or more) systems of belief are not allowed to come into rational contact with each other. But here, the formerly unconscious belief *is* brought into rational contact with the rest of the person's conscious thought even while being experienced as her belief, and it is rejected as false. That is precisely how it is integrated into her total subjective perspective: as an aspect of that perspective that is false. This integration is assisted by the way in which the belief is rendered intelligible through an experientially based, emotionally rich understanding of the dynamic sources of the belief—an understanding that arises through conscious experience of the fears and wishes that hold it in place, so that they too, as well as their relation to the belief, all become part of the single subjective perspective. To come to understand—on the basis of a rich conscious experience of the relevant emotions—that one's belief is, for instance, an infantile defensive reaction to infantile fears is (p. 318) simultaneously to render it intelligible, to accept it as what it is, and to put it in its place—even if it continues to be an ineliminable part of one's total perspective.

Is There a Superior Alternative Description?

I have been arguing that we have to allow for the possibility of self-consciously believing something that one recognizes to be false, if we are to make sense of the very idea of integrating the mind by bringing unconscious beliefs into consciousness. One possible response is to feel that my argument just can't be right: since (it is said) we can make no sense of self-consciously believing something one knows to be false, we must have been wrong to take this as a matter of belief to begin with.

In a well-known paper, the philosopher Tamar Gendler (2008) urges the explanatory importance of what she calls 'aliefs': automatic, associative, arational clusters of representational content, affect, and associated behavioural routines. It might be suggested that what I have been talking about are really *belief-discordant aliefs* that are brought into consciousness, not fully fledged beliefs that would generate a Moore-paradoxical position of the sort I have been describing. It might similarly be said that in many of these cases it would be more natural to express the attitude using *affective* rather than *cognitive* language, saying something like 'I feel like I've stolen everything good in my life, but of course I know that's not so'.[12]

In some cases, either of these forms of description might be apt. In particular, it might be the case that many of what psychoanalysts call 'unconscious beliefs' consist in associative clusters of affective and behavioural dispositions. From a clinical perspective, however, the important point is this. Even if they were to begin life as 'aliefs', dynamically unconscious beliefs can undergo a *developmental trajectory* in which they acquire determinate propositional content and are brought into relation with the rest of the person's conscious view of the world in ways that are emotionally and motivationally

transformative. It is precisely this work that is amongst the goals of the analytic therapy. But the result might well be a conscious attitude that conflicts with the person's other conscious beliefs.

Likewise, casting these phenomena in broadly affective terms is fine, so long as one does not lose track of the crucial fact that the attitudes in question *conflict with* one's settled view. So long as one keeps this point clearly in mind, the terminology does not matter. Where the terminology can become important is in the clinical setting, since talk of what one 'feels' can be used to distance oneself from acknowledging how one actually takes things to be. Gardner's characterization of defence is apt here: 'an operation on mental content that represents the cause of anxiety in such a way as to reduce or (p. 319) eliminate anxiety' (1993: 145). Talk of what one 'feels' can function defensively to avoid the experience of psychic conflict by disguising unwanted or unacceptable aspects of one's view of what is the case.

It would likewise be a mistake to think that what I have been describing is best understood as a situation in which it merely *seems* or *appears* to the person as if things are a certain way. For her to deny that she holds the belief and to go straight to 'it merely seems or appears to me as if I have stolen everything good', would be for her to deny something crucial. When the belief was functioning unconsciously, this *was* how things were for her, though she didn't recognize it. And she is attached to this way of viewing her world, insofar as doing so is fuelled by and satisfies particular motivations and defensive responses to them. Believing that things are this way thus does important psychic work for her. None of that necessarily changes as these matters come into consciousness. Thus for her to say that this is not belief (but rather a mere 'seeming' or 'appearance') would be for her to evade these truths about herself in a way that enables the processes that maintain the belief to continue unmodified and unexamined. For the same reason, to characterize her state as mere 'seeming' or 'appearance' would miss— and render unintelligible—the clinical fact, emphasized by Britton, that once a dynamically unconscious belief becomes conscious, a process of *mourning* is often required before the belief can be given up (1995: 22).

Similar points apply to the thoughts that the person merely has an 'inclination' to believe —but doesn't actually hold the belief—or that only 'a part' of her believes. To go straight to such self-characterizations might very well be a motivated attempt to evade the difficult conflicts presented by the states in question or to avoid approaching the motivations that hold the belief in place. This would enable one to maintain the conflicting view of things at a deeper level even while consciously avowing it in a mischaracterized form. Such a self-characterization would short-circuit the difficulties involved in integrating the false unconscious belief. Psychoanalytic-sounding ideas, too, can be used for defensive ends.[13]

Moore-paradoxical belief attributions also cannot be avoided by emphasizing implicit, non-propositional procedural schemata for interpersonal relating. All psychoanalytic writers in this vein grant that there is a place clinically for the analysand's conscious

articulation of implicit models for relating.[14] This cannot be accomplished except by the analysand's acknowledging from a first-person position how he takes things to be in his interaction with the analyst (e.g. 'I can't reach out to you for comfort, because you (p. 320) disapprove of comfort-seeking').[15] Since what is being articulated is how things *are* from the patient's subjective viewpoint, this amounts to the conscious expression of belief. And there is no way for the patient to fully own—from the inside, as it were—the full pattern of inhibition, motivation, emotional response, silence, etc. that is involved here, unless he recognizes this as his perspective. But in some cases this will involve a proposition which he regards as false.

An additional point is relevant here. Theorists who emphasize implicit procedural schemata do not hold that this theoretical construct can entirely replace the clinical role played by the concept of unconscious belief and its articulation in the process of working through, since these schemata can generate distinct psychodynamically unconscious beliefs. Fonagy, for instance, writes that 'Consciousness of the beliefs generated by such implicit memory models is crucial if the patient is to acquire the power to inhibit or modify them through the creation of a second-order representation of their inner experience . . .' (Fonagy 1999: 219). Here again, a puzzle about the integration of false unconscious belief will arise.[16]

Someone who thinks that the self-conscious self-attribution of false belief is just *impossible* might instead suggest that as the belief comes into consciousness, the analysand oscillates between incompatible, internally coherent viewpoints in order to find a resting place in one and give up the other. But this picture is not adequate either. First, there is often no larger viewpoint on the world available to the person within which the formerly unconscious belief could be made to look rational. Because the belief is sustained not by rational relations but rather by a tissue of phantasy, wish-fulfilment, and defensive processes, to give it a place in a larger rationally coherent view would require the wholescale creation of a massively delusive view of the world. Second, this picture is overly optimistic about the extent to which it is possible to rid oneself of such beliefs. Often they remain, and they need to be given a place within one's larger subjective perspective if one is to have hope of putting together a stable life. The picture of 'oscillation' cannot allow for this. Finally, to think that this is how things should *always* go is to maintain a fantasy of being able to get rid of psychic materials one does not want. It would be better in many cases to aspire to fully occupy one's subjective perspective with acceptance and understanding, where that very perspective includes a well-developed, emotionally rich, and motivationally effective conception of its own weaknesses and deficiencies.

I think that this conclusion is inevitable for any psychodynamic therapeutic approach that takes the patient's total first-person, subjective perspective seriously as a realm for exploration, articulation, and productive elaboration through speech and avowal.

(p. 321) What Integration Might Look Like

Various positive endpoints are possible. With work, the person may give up the belief and correctly understand herself as subject to an unshakeable emotionally resonant 'appearance' or 'seeming' arising from understood psychodynamic forces. Through underlying psychological shifts she might lose even this 'seeming' or 'appearance'. In other cases, the underlying motivations might remain but gain their satisfaction not in the person's cognitive life, not in the realm of seemings and belief, but rather in the realm of treasured conscious fantasy and play. However, in many cases the endpoint is one in which despite what the person might want, it would be untrue for her to deny that she believes that she is a thief. In these cases the false belief is consciously maintained but given its place within a broader perspective that acknowledges its falsehood, renders it intelligible, and ameliorates its effects.

Even here there can be variation along multiple dimensions. One such dimension is the precise *way* in which the person now occupies the subjective standpoint of the formerly unconscious belief.

To see this, consider that the ability to use 'I' in expressive self-ascription of the attitude doesn't yet guarantee incorporation of this belief into the right sort of larger self-conscious perspective. After all, someone could, at a given moment, consciously occupy the subjective position of the belief that he is a thief and say, in a melancholic, chastened tone, 'I'm an awful person; I deceive people in order to get their love. I'm nothing but a thief.' He might continue, in response to a therapist's comment, 'You believe that you are a thief', 'Yes, I do. I steal from everyone.' This person has not yet integrated his formerly unconscious false belief, because he is now *in the grip of it*. It isn't yet properly incorporated into the right sort of broader perspective. (This position might nonetheless be an important stage in a developmental process, and someone engaged in 'working through' may slip in and out of it.)

Someone might instead be able to self-attribute the belief from the subject position, consciously having the thoughts, feelings, and motivations that go along with it, and yet battle against it in light of its falsehood. He might fight and struggle to resist some of the dispositions it involves, for instance by effortfully correcting false interpretations it generates in particular interactions. He might be able to place all of this in the context of a broader understanding of the psychodynamic forces at issue and an ability to recognize them as they appear and function in situ. In these ways, he *manages* the false belief.

This situation is a step towards psychic integration, but precisely in its aspect of intrapsychic struggle, it threatens to break apart. To put it metaphorically, the belief that he is a thief is pressing to bust loose, and if it breaks free he might oscillate back towards a position in which he is in the grip of the belief. The person's overall mental state is thus *organized* at the moment, but it lacks a measure of *stability*. Like a peace imposed by armed garrison, it is subject to internal pressures that threaten to undo it.

A notable feature of this last case is that there is a sense in which the person restrains himself from fully occupying the subjective position of the belief, precisely because of the threat it is felt to pose. For instance, he might not be able to give voice to this belief by saying 'I am a thief' without sliding back towards being in its grip. While he can wholeheartedly self-ascribe the belief, he might feel enormous anxiety at thinking the first-order thought, 'I am a thief'. That belief's subjective perspective thus hasn't yet been stably and fully incorporated into a larger, unified perspective in which he simultaneously occupies the standpoint of the belief even while recognizing its falsity.

How could he give some form of first-order expression to the belief even while maintaining his hold on its falsehood? Imagine that while he recognizes that his belief is false, he is no longer engaged in intra-psychic struggle against it. Instead, he is able to say, with warm, ironic good humour and a twinkle in his eye, 'Of course I am a thief. After all, I'm doing it right now! I've gotten you to pay undivided attention to me for a full hour, and I've given you no love or attention in return.' Here, tone is everything. We have to imagine that the person is both giving voice to his belief that he is a thief and also *in the fact that he is speaking with ironic good humour* giving voice to what he would otherwise express by saying 'Of course, I'm not a thief'. He thus gives voice to both sides of the conflict while putting each in its proper place. And this single utterance doesn't just express the conflict; it expresses his total position regarding both the belief and the conflict it is involved in. For instance, his utterance also expresses, with good-humoured irony, the rationalizing pressure to seek confirming evidence for the belief that he is a thief ('See, I'm doing it right now!'), and the gentle, self-accepting, good humour with which all of this is said expresses his recognition that this belief (with all that it involves) is an immature solution to an immature problem. What we see here is thus an expression of the full, stable incorporation of the false belief into a single, unified self-conscious perspective.[17]

This person can give voice to both sides of the conflict without any oscillation in his total position—that is, without any change in his total mental state, including its associated dispositions and patterns of thought, feeling, and reaction. The utterance exhibits an apt, settled pattern of relations between both sides of the conflict and also an apt, settled orientation towards the conflict itself. The person thus has the ability to speak with one voice—a voice that expresses the totality of his complex overall perspective all at once. Indeed, it is this ability that is a key reason for describing this as *one* contemporaneous complex subjective perspective comprising both the subjective position of the belief and recognition of its falsehood. What underwrites such an ability are the functional relations amongst the person's beliefs, thoughts, feelings, motivational states, and ongoing patterns of thought, feeling, motivation, and reaction, including the relatively stable maintenance of an organization over time that includes an overall good grip on reality.

This, then, is one way things can look when a person has successfully integrated a false unconscious belief. It is not the only way, and it may be but a stage in a larger process. (p. 323) There is still the question of what motivations lie behind this belief. Moreover, though this resolution is well attuned to reality and stable so far as the particular conflict

is concerned, it may be threatened or undercut by other psychodynamic processes or co-opted for other purposes. Moreover, certain defensive structures can look superficially similar to this resolution; here the difference might be marked, for instance, by the superficiality of the irony, the lack of warm self-acceptance, the casual flimsiness of the patient's claim to know that he is not a thief, or the panic-concealing giddiness of his statement that he is one.

My use of the phrase 'speaking with one voice' emphasizes a connection with an important theme in recent work by Jonathan Lear. Lear has articulated a broadly neo-Aristotelian concept of integration involving excellent communication between the rational and non-rational aspects of the psyche, communication that enables a person to move forward freely to satisfy both the person's conscious aims and unconscious yearnings (Lear 2017). As he puts it:

> If the voices of the non-rational soul are an occasion for a creative, in-tune and thoughtful response from reason; and if, in turn, reason is able to enliven and free up the voices of the non-rational soul, as it channels them into a life worth living, we can give content to the thought that this is a rich form of speaking with the same voice.

> (Lear 2017: 49)

Here Lear calls attention to the way in which two systems or 'parts' of the psyche can work in harmony despite their different modes of functioning: they are, at bottom, in a certain kind of accord. This is an important form of integration. However, it differs from the mode of integration which I have been highlighting: the integration of conflicting mental states into a whole that neither evades nor elides the conflict but rather incorporates it into a broader, well-functioning, self-conscious standpoint.

Lear's own clinical examples do not thematize the relation between these two modes of integration. One of his clinical examples—the case of Mr B—is described in such a way that it appears that excellent communication is achieved without the patient's consciously integrating whatever conflict is holding him back (2017: 42). In the example of Ms A, by contrast, the patient self-consciously experiences the conflict for what it is, and this appears to play a role in enabling her to make a request that had been difficult for her (2017: 22). This is a moment that is comparable to what I have described in the integration of unconscious belief.

It is an open question for me to what extent and in what ways excellent communication between the rational and non-rational aspects of the psyche involves consciously experiencing and integrating conflicting mental states, just as it is an open question in what way(s) the latter form of integration might depend upon the form of good communication that Lear describes. The point that I want to mark for now is just this: because these phenomena are conceptually distinguishable, these questions need to be

considered. As Lear has helped us see, we shouldn't assume in advance that we know what integration is or should be.

(p. 324) Philosophical Implications: Psychic Unity, the First-Person Perspective, Transparency, and Rational Self-Constitution

I have argued that if we are to make sense of the possibility that bringing unconscious beliefs to consciousness could be a path to psychic health, we must make sense of the idea of self-consciously occupying the standpoint of a belief that one knows to be false. I have tried to show how to do so. In closing, I will briefly sketch some important implications for several large issues in philosophy.

a. There is a strong tendency amongst philosophers to view both the unity of the subject and psychic integration as matters of *rational* unity. For instance, in his defence of the broadly Freudian idea that the mind can contain partitions between distinct psychic systems, Donald Davidson comments, 'I postulate such a boundary somewhere between any (obviously) conflicting beliefs.... Such boundaries … are conceptual aids to the coherent description of genuine irrationalities' (Davidson 1986: 211; see also 1982). Such aids are needed, Davidson tells us, because 'two obviously opposed beliefs could coexist only if they were somehow kept separate, not allowed to be contemplated in a single glance': these are 'beliefs which, allowed into consciousness together, would destroy at least one' (Davidson 1997: 220). But the result of the partition, he says, is 'a single mind not fully integrated' (1997: 221).

These formulations presuppose a picture of psychic integration as *rational* integration, resulting in the rational abandonment of one or the other of the conflicting beliefs. The integrated mind is thus free from obvious conflict, on this picture. However, we have seen that self-conscious belief isn't tied that closely to rationality, nor does psychic integration necessarily involve the removal of rational conflict.[18]

(p. 325) Sydney Shoemaker deploys framing assumptions similar to Davidson's in considering Moore's paradox. He writes, of a person who affirms 'I believe that p but p is false', that 'the only way to save the coherence of this case is to suppose that it involves there being a divided mind. One part of a person's mind believes something, and another does not, and the part that does not believe it ascribes the belief to the part that does' (Shoemaker 2012: 249). Here Shoemaker gets things exactly backwards. The analysand's ability to sincerely say or think such a thing is a manifestation precisely of the overcoming of psychic division: it's from one and the same perspective—that of the conscious subject of the belief in question—that the analysand both self-ascribes the belief and acknowledges its falsehood.

At bottom, the lesson is this. There is a widespread view that it cannot possibly be correct to ascribe an obviously irrational combination of contemporaneous conscious attitudes to a person. Consideration of the clinical task of integrating false unconscious beliefs shows that this view is simply incorrect. The unity of a single, self-conscious, subjective perspective need not be a rational unity.[19]

b. I turn now to an issue prominent in recent philosophical discussions of knowledge of one's own beliefs and other attitudes. It is a familiar point in these discussions that there is a crucial difference between ordinary first-personal awareness of our emotions, beliefs, desires, and other attitudes, on the one hand, and, on the other hand, knowledge we might gain about these matters in other ways. For instance, someone might learn of her unconscious belief by accepting an interpretation offered by her analyst or by drawing an inference from indications in her behaviour and conscious thought and feelings. She doesn't thereby gain the sort of knowledge that we ordinarily have of our own beliefs. The difference is that her self-attribution is analogous to a third-personal psychological attribution: she does not self-ascribe the belief from the standpoint of the conscious subject of the very belief being self-ascribed. As it is put in the literature, she does not stand in a genuinely *first-personal* relation to her belief (Moran 2001).

According to a long and prominent philosophical tradition, the genuinely first-personal relation to one's attitudes is inextricably linked to considerations of rationality and a conception of the self as rational agent. Perhaps the most prominent proponent of such a view is Kant, who writes that 'it is only as intelligence that [a person] is his proper self' and that if certain desires cannot be viewed by the individual as the products of reason, 'He does not even hold himself responsible for these inclinations or impulses or attribute them to his proper self' (Kant *Grundlegung* 457–8; Beck's translation p. 7). Themes along these lines have shaped the thinking of a large number of contemporary philosophers, not all of whom identify as specifically Kantian (Bilgrami 2006; Boyle 2009; Moran 2001; Rödl 2007). For instance, some hold that self-attribution of beliefs from the first-person stance necessarily involves a stance of deliberative agency and rational endorsement. Others hold that in first-personal attributions of belief, one sees the belief as the product of one's rational faculties.

(p. 326) The points I have been urging show such views to be mistaken. Self-attribution from a first-person stance requires consciously occupying the attitude's subjective perspective on the world. In the cases I have been discussing this does not require or allow for rational endorsement of the belief or even seeing it as the product of the proper functioning of one's rational faculties. Nor in these cases does the first-person relation involve taking a stance of deliberative agency towards the belief, since one might very well recognize that one can't get rid of the belief through rational deliberation. Rather, in these cases a genuinely first-personal stance involves seeing the world from the subjective standpoint of the belief even while regarding it as false and as the product of non-rational belief-forming processes involving wish, phantasy, and defence. What is required here is acknowledgement and acceptance of the reality of one's psychic life, building on a certain kind of emotional openness to that which one finds repugnant or

unreasonable.[20] This is not reducible to any sort of intellectual exercise, nor is it compatible with identifying oneself (as subject) with one's rational capacities in particular.

Still, something is right in thinking that our reasoning capacities, and our understanding of ourselves as possessing and aiming to exercise such capacities, are relevant to our ability to take up a first-personal relation to our beliefs. Our ability to acknowledge and accept our false beliefs depends fundamentally on our capacities for rational reflection and response to reality. Without those capacities, we could not contain the conscious false belief within a broader conscious perspective that puts it in its place, but would rather be the helpless victims of unresolved conflict.

c. Problems also arise for the related accounts of first-personal self-knowledge which take 'transparency' to be central. Gareth Evans claimed that in self-ascribing a belief about whether there will be a third world war 'I must attend … to precisely the same outward phenomena as I would attend to if I were answering the question' whether there will be a third world war (1982: 225): that is, to determine whether I believe that p, I ask whether p and then if I find that p is true (or false), I append 'I believe that . . .' to that result. As Moran puts it, '[T]he claim [of the Transparency Condition] … is that a 1st-person present tense question about one's belief is [to be] answered by reference to (or consideration of) the same reasons that would justify an answer to the corresponding question about the world' (Moran 2001: 62). On this conception, present-tense self-attribution of belief from the first-person standpoint is dictated by rational consideration of the relevant evidence and facts concerning the matters the belief is about.

This view cannot make sense of the clinical phenomena. When an analysand integrates a formerly unconscious belief that is plainly false, she cannot self-attribute the belief by following the transparency procedure; after all, she recognizes that the belief is false, so when she asks whether p her answer will point away from self-attribution of belief that p —at least if she follows the transparency procedure. Still, she stands in a properly first-personal relation to her belief that p nonetheless. Her relation to that belief is not merely 'attributional' or 'theoretical'; she occupies the subjective perspective of the belief and (p. 327) self-attributes the belief from that position. Properly first-personal self-attribution thus need not conform to the transparency condition.

This point is of clinical importance. Precisely because the transparency procedure ties self-attribution to one's judgement about what is the case or what the relevant reasons support, to take up the transparency approach in the clinical situation would be to refuse to admit into one's self-conscious view any belief that one can see fails to fit with the relevant evidence or facts. The stance of Transparency would thus be a manifestation of resistance to the analytic process or a defence against conscious psychic conflict in these cases. In the clinical setting, the key question is, 'How am I oriented towards myself and my world?' If one is tempted to answer that question by focusing on the question, 'What is there reason for here?', one might well be led into a defensive, false self-conception (Bell and Leite 2016; Lear 2011: 56).

Some accounts of self-knowledge emphasize transparency without framing it in terms of the Kantian themes highlighted earlier. Clinical cases raise a challenge for these views as well.

For instance, according to Alex Byrne's (2005, 2011) 'inferentialist' account, first-personal awareness of our own beliefs is secured through an inference in accordance with the 'Doxastic Principle', 'If p, then conclude that you believe that p'. Byrne emphasizes that this would be a distinctively first-personal way of acquiring such knowledge; the inference's reliability depends precisely on the fact that it concerns *one's own* beliefs. However, the patient who comes to see 'from the inside' that she believes—quite incorrectly—that she has stolen everything good in life, has genuinely first-personal knowledge of this belief, and yet we cannot reconstruct this knowledge in terms of the reasoning Byrne proposes. First, the person would not reason in the way he suggests ('I have stolen everything good in my life, so I believe that I have stolen everything good in my life'), because she regards the premise as false. Second, it is hard to see how an inference from a premise a person correctly regards as false could give her knowledge of the conclusion. Her first-personal awareness of this belief consequently cannot be grounded in any such reasoning.

d. I turn to one last issue: the constitution of the self. A prominent strand of thought in contemporary philosophy urges that *my endorsement* is what makes the difference between that which belongs to 'me' as subject and that in my psychology to which I relate only as an object of description, report, and possibly management (Frankfurt 1988; Korsgaard 2009).[21] On this view my unity as a subject is threatened to the extent that I have not yet sorted conflicting elements into 'me' and 'not-me' through deliberation and endorsement or rejection. Korsgaard gives characteristically sharp expression to the point.

> Because human beings are self-conscious, we are conscious of threats to our psychic unity or integrity. Sometimes these threats spring from our own desires and impulses … we deliberate in the face of threats to our integrity, and as against them … we must repress them in order to be one, to be unified, to be whole (2009: 26).

(p. 328) Korsgaard does not merely mean that we give order to our psychologies by choosing to act on some motives and beliefs but not others. Rather, she thinks of deliberation and choice as forging a distinction between what is genuinely 'us' and what is not. 'That's what deliberation is: an attempt to reunite yourself behind some set of movements that will count as your own' (Korsgaard 2009: 213); through deliberation and choice we engage in 'the endorsement of our identities, our self-constitution' (2009: 43). So, for Korsgaard self-unity is achieved through extrusion: I cannot recognize from the first-person point of view that something belongs to me *as subject* which I do not and cannot endorse, but rather must regard it merely as 'a product of some force that is at work *on* me or *in* me' (2009: 18–19).

This gets things backwards from the clinical point of view. What is true, given a psychodynamic framework, is rather that psychic health requires us to *integrate* these threats 'in order to be one, to be unified, to be whole.' That doesn't mean that we must endorse them or choose to act upon them. It means that we must bring them within our purview as self-conscious subjects, understanding them as *us*, as part of our identity, as part of who we are.[22]

References

Arlow, J. (2008/1969). 'Unconscious fantasy and disturbances of conscious experience'. *Psychoanalytic Quarterly* 77(1): 21–46.

Auchincloss, E. L. and Samberg, E. (eds) (2012a). 'Fantasy'. In *Psychoanalytic Terms and Concepts*. New Haven, CT: Yale University Press.

Auchincloss, E. L. and Samberg, E. (eds) (2012b). 'Splitting'. In *Psychoanalytic Terms and Concepts*. New Haven, CT: Yale University Press.

Beebe, B. and Lachmann, F. (1994). 'Representation and internalization in infancy: Three principles of salience'. *Psychoanalytic Psychology 11*: 127–165.

Bell, D. and Leite, A. (2016). 'Experiential self-understanding'. *The International Journal of Psychoanalysis 97*(2): 305–332.

Bilgrami, A. (2006). *Self-Knowledge and Resentment*. Cambridge, MA: Harvard University Press.

Bohleber, W., Jiménez, J. P., Scarfone, D., Varvin, S., and Zysman, S. (2015). 'Unconscious phantasy and its conceptualizations: an attempt at conceptual integration'. *International Journal of Psychoanalysis 96*(3): 705–730.

Boyle, M. (2009). 'Two kinds of self-knowledge'. *Philosophy and Phenomenological Research 78*(1): 133–164.

(p. 329) Britton, R. (1995). 'Psychic reality and unconscious belief'. *The International Journal of Psychoanalysis 76*(1): 19–23.

Britton, R. (1998). *Belief and Imagination*. New York, NY: Routledge.

Byrne, A. (2005). 'Introspection'. *Philosophical Topics 33*: 79–104.

Byrne, A. (2011). 'Transparency, belief, intention'. *Proceedings of the Aristotelian Society* Supplementary Volume *85*: 201–221.

Child, W. (1996). *Causality, Interpretation, and the Mind*. Oxford: Oxford University Press.

Coliva, A. (2015). 'How to commit Moore's Paradox'. *Journal of Philosophy 112*(4): 169–192.

Davidson, D. (1970/1980). 'Mental acts'. In *Essays on Actions and Events* (pp. 207–225). Oxford: Oxford University Press.

Davidson, D. (1982/2004). 'Paradoxes of irrationality'. In *Problems of Rationality* (pp. 169–187). Oxford: Oxford University Press.

Davidson, D. (1983/2001). 'A coherence theory of truth and knowledge'. In *Subjective, Intersubjective, Objective* (pp. 137–153). Oxford: Oxford University Press.

Davidson, D. (1984). *Essays on Truth and Interpretation*. Oxford: Oxford University Press.

Davidson, D. (1986/2004). 'Deception and division'. In *Problems of Rationality* (pp. 199–212). Oxford: Oxford University Press.

Davidson, D. (1997/2004). 'Who is fooled?' In *Problems of Rationality* (pp. 213–230). Oxford: Oxford University Press.

Eagle, M. (2013). 'The implications of conceptual critiques and empirical research on unconscious processes for psychoanalytic theory'. *Psychoanalytic Review 100*(6): 881–917.

Erreich, A. (2003). 'A modest proposal: (re)defining unconscious fantasy'. *Psychoanalytic Quarterly 72*(3): 541–574.

Evans, G. (1982). *Varieties of Reference*. Oxford: Oxford University Press.

Feldman, M. (2007). 'Addressing parts of the self'. *International Journal of Psychoanalysis 88*(2): 371–386.

Fonagy, P. (1999). 'Memory and therapeutic action'. *International Journal of Psychoanalysis 80*(2): 215–223.

Frankfurt, H. (1988). 'Freedom of the will and the concept of a person'. In *The Importance of What We Care About* (pp. 11–25). Cambridge: Cambridge University Press.

Freud, S. (1915). 'The unconscious'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud,* Vol. 14 (pp. 159–215). London: Hogarth Press.

Freud, S. (1917). *Introductory Lectures on Psychoanalysis*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud,* Vol. 15. London: Hogarth Press.

Freud, S. (1927). 'Fetishism'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 21 (pp. 152–157). London: Hogarth Press.

Freud, S. (1940a). 'An outline of psychoanalysis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 23 (pp. 144–208). London: Hogarth Press.

Freud, S. (1940b). 'Splitting of the ego in the process of defence'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud,* Vol. 23 (pp. 275–278). London: Hogarth Press.

Gardner, S. (1993). *Irrationality and the Philosophy of Psychoanalysis.* Cambridge: Cambridge University Press.

Gendler, T. S. (2008). 'Alief and belief'. *Journal of Philosophy 105*(10): 634–663.

(p. 330) Gertler, B. (2011). 'Self-knowledge and the transparency of belief'. In Anthony Hatzimoysis (ed.), *Self-Knowledge* (pp. 125–145). Oxford: Oxford University Press.

Heal, J. (1994). 'Moore's paradox: a Wittgensteinian approach'. *Mind 103*(409): 5–24.

Hursthouse, R. (1991). 'Arational actions'. *Journal of Philosophy 88*(2): 57–68.

Isaacs, S. (1948). 'The nature and function of phantasy'. *International Journal of Psychoanalysis 29*: 73–97.

Kant, I. (1990). *Foundations of the Metaphysics of Morals*, trans. L. W. Beck. New York: Macmillan.

Kernberg, O. (1966). 'Structural derivatives of object relationships'. *International Journal of Psychoanalysis 47*: 236–252.

Klein, M. (1946). 'Some notes on some schizoid mechanisms'. In *Collected Works of Melanie Klein*, Vol. 3: *Envy and Gratitude and Other Works* (pp. 1–24). London: Hogarth Press.

Korsgaard, C. (2009). *Self-Constitution: Agency, Identity, and Integrity*. Oxford: Oxford University Press.

Lear, J. (2003). *Therapeutic Action*. New York: Other Press.

Lear, J. (2011). *A Case for Irony*. Cambridge, MA: Harvard University Press.

Lear, J. (2017). *Wisdom Won From Illness*. Cambridge, MA: Harvard University Press.

Leite, A. (2016). 'Changing one's mind'. *Philosophy and Phenomenological Research*, version of record online 7 Nov. 2016. doi:10.1111/phpr.12332

Leite, A. (2018). 'Expressivism, psychic health, and the first-person stance'. *Journal of Personality Disorder* (special issue on self-knowledge and personality disorder) *32*(3): 351–373.

Lewis, D. (1982). 'Logic for equivocators'. Noûs *16*(3): 431–441.

Lyons-Ruth, K. (1999). 'The two-person unconscious: Intersubjective dialogue, implicit relational knowing, and the articulation of meaning'. *Psychoanalytic Inquiry 19*: 567–617.

Mölder, B. (2010). *Mind Ascribed: An Elaboration and Defense of Interpretivism*. Philadelphia, PA: John Benjamins.

Moore, G. E. (1993). 'Moore's Paradox'. In T. Baldwin (ed.), *G. E. Moore: Selected Writings* (pp. 207–212). London: Routledge.

Moran, R. (2001). *Authority and Estrangement*. Princeton, NJ: Princeton University Press.

Parrott, M. (2015). 'Expressing first-person authority'. *Philosophical Studies 172*: 2215–2237.

Rödl, S. (2007). *Self-Consciousness*. Cambridge, MA: Harvard University Press.

Schwitzgebel, E. (2002). 'A phenomenal dispositional account of belief'. *Noûs 36*(2): 249–275.

Shoemaker, S. (2012). 'Self-intimation and second-order belief'. In D. Smithies and D. Stoljar (eds), *Introspection and Consciousness* (pp. 239–258). Oxford: Oxford University Press.

Steiner, J. (1993). *Psychic Retreats: Pathological Organizations in Psychotic, Neurotic and Borderline Patients*. London: Routledge.

Stern, D. N., Sander, L. W., Nahum, J. P., Harrison, A. M., Lyons-Ruth, K., Morgan, A. C., Bruschweilerstern, N., and Tronick, E. Z. (1998). 'Non-interpretive mechanisms in psychoanalytic therapy: the "something more" than interpretation'. *International Journal of Psychoanalysis 79*: 903–921.

Williams, P. (1996). 'Panel report—psychic reality: theoretical concepts'. *International Journal of Psychoanalysis 77*: 79–84.

Wittgenstein, L. (2001/1953). *Philosophical Investigations*. Oxford: Blackwell Publishing.

Wollheim, R. (2003). 'On the Freudian unconscious'. *Proceedings and Addresses of the American Philosophical Association 77*(2): 23–35.

Notes:

(1) Schwitzgebel (2002) offers a broadly dispositional approach to belief that fits well with the perspective sketched here.

(2) There are two options here because someone could have patterns and dispositions of conscious response—and be aware of the world in ways—that constitute occupying the subjective perspective of the attitude, and yet be unable to self-ascribe it. For instance, she might lack facility with relevant mental state concepts. More subtly, to consider just one sort of alternative in which she possesses facility with the relevant concepts, she

might consciously experience reactions of anger or envy but for defensive reasons be unable to experience them *as* anger or as envy. Finkelstein (this volume) offers an expressivist account of the conscious/unconscious distinction which attempts to cover all of the relevant cases. However, it is doubtful that his account adequately handles clinical examples in which a person directly expresses her attitude through self-attribution and yet fails to understand in the requisite way what she is doing (Bell and Leite 2016; Leite 2018). More complicated versions of such examples cause trouble for Finkelstein's account even when it is supplemented with the requirement (Finkelstein, this volume) that the person be able to 'co-expressively gloss' the expression of the attitude.

(^{3}) Only a great deal of experience with the person will enable a belief attribution of this specificity. For instance, some of the data I have described is compatible with an attribution of the belief 'I am an imposter', but someone who thinks himself an imposter rather than a thief will exhibit a different profile of thoughts and emotional responses. However, some measure of indeterminacy may be an inevitable part of mental state attributions as well. (Nothing I have said here addresses questions of clinical technique such as how the analyst should regard her own conjectures about the patient or how she should bring them into the clinical exchange.)

(^{4}) This is a schematic characterization of the mechanism of 'disavowal' proposed by Freud in his attempt to account for certain cases of fetishism (1927, 1940a: 201–4, 1940b).

(^{5}) The Kleinian tradition uses the spelling 'phantasy' to mark off the specifically psychoanalytic concept, as do many other analysts writing in Britain. However, many non-Kleinian psychoanalysts, especially those in America, prefer the spelling 'fantasy'. For convenience, I will follow the Kleinian spelling. No tendentious theoretical assumptions are thereby intended.

(^{6}) This difficult task often isn't helped by saying that a part of the patient hates the object while another part loves the object. As Michael Feldman puts it, 'while the patient might be able to understand such interpretations intellectually, [she] may not be able to use it in an insightful way, to promote psychic integration' (2007: 383).

(^{7}) Not all cases are of this sort; for instance, many repressed beliefs in relation to trauma are not false.

(^{8}) I don't deny that sometimes we simply don't know what to say about what the person believes or knows. However, the case relevant here isn't like that. It is one in which the person has the dispositions (described earlier) involved in unconsciously believing that he is a thief and *also* has the dispositions involved in believing—indeed, knowing—that he isn't one. The complex interplay between these dispositions explains the variegated pattern of actual behaviour and reactions.

(^{9}) The integration of unconscious belief thus poses a particularly sharp challenge for this philosophical tradition. It has been noted in recent work on Moore's paradox that in

situations of irrational or superstitious belief, a person might recognize that she has a belief even while acknowledging its falsity (Gertler 2011; discussion in Coliva 2015). However, it hasn't been clear in these examples whether the person occupies the subjective position of the belief at the time of the self-attribution. If not, then such examples are beside the point. Just as it is no surprise that a person could recognize on the basis of her own behaviour that she has a false unconscious belief, it is no surprise that self-attribution is straightforward when one is alienated from one's belief and does not at that moment occupy its subjective standpoint on the world. This is why the integration of unconscious false belief is a key test case.

(^{10}) This phenomenology is distinct from *what* is presented as being the case, but rather concerns the manner in which it is so presented.

(^{11}) It is sometimes asserted that understanding belief as involving 'commitment'—say, a 'commitment' to the truth of the believed proposition—precludes intelligible self-ascription of false belief (Coliva 2015). However, on any ordinary conception of commitment you can perfectly well recognize that you have and are expressing a commitment that you shouldn't have: a commitment that is erroneous, mistaken, or misguided. Admittedly, 'commitment' could be stipulatively defined in such a way that 'commitment' to the truth of p is incompatible with the self-attribution of false belief. However, one had better not understand conscious belief in such terms. Such an account has the consequence that self-conscious, deliberative change of mind is impossible—which it surely isn't (Leite 2016).

(^{12}) Thanks here to Sarah Majid.

(^{13}) Another sort of short circuit would arise if the analysand defensively rolled up his sleeves and got to work on trying to change the belief before truly coming to occupy it (Wollheim 2003).

(^{14}) Fonagy, for instance, writes: 'Therapeutic action lies in the conscious elaboration of preconscious relationship representations' (1999: 218). Stern et al. comment: 'In the course of an analysis some of the implicit relational knowledge will get slowly and painstakingly transcribed into conscious explicit knowledge' (1998: 918). Lyons-Ruth writes: 'If relational knowing is as much implicit and procedural as symbolic, the work of elaborating new implicit procedures for being with others must occur at enactive *as well as* symbolic levels … For an adult patient, more collaborative and inclusive dialogue may involve partially translating previously implicit procedural knowing into words' (1999: 608, 611, italics added).

(^{15}) This example is suggested by a discussion in Lyons-Ruth (1999: 596).

(^{16}) Even Stern et al. recognize the clinical importance—and difficulty—of bringing dynamically unconscious beliefs to consciousness. They write: 'The process of rendering repressed knowledge conscious is quite different from that of rendering implicit knowing

conscious. They require different conceptualisations. They may also require different clinical procedures, which has important technical implications' (1998: 918).

(¹⁷) There is a link here—which only became clear to me after writing this chapter—to Lear's important work on irony and therapeutic action (2003: esp. 118–33; see also Lear 2011).

(¹⁸) Psychoanalysts, too, are sometimes tempted by the conception of integration as an absence of conflict. For instance, Morris Eagle writes, 'In his later writings, [Freud] writes that in health there is a "seamlessness" between the id and the ego, and also that the overarching goal of psychoanalysis is unity of the personality. In taking this general perspective, Freud aligns himself with those whose philosophical and spiritual ideal is being at one with oneself, a man or woman of one piece. These figures include Confucius … who reported that at the age of seventy … the dictates of his heart and his sense of right and wrong were now one and the same, and Kierkegaard … who wrote that purity of heart is to will one thing. The primary value underlying this idea is … the experience of being at peace with oneself…. One way of putting it is to say that psychoanalysis moves a bit out of what appeared to be the near-exclusive orbit of the Enlightenment vision and closer to philosophical traditions characterized by a central concern with the achievement of inner harmony' (Eagle 2013: 908). Here we see a slide from 'seamlessness' (which is a functional/structural concept) to an ideal of integration as the removal of conflict. However, seamlessness and the unity of the personality (understood in functional/structural terms) do not in fact require the removal of conflict, but rather acceptance of the conflict—of both sides of the conflict as 'me'—and the discovery or creation of a way to move forward given the fact of conflict. This alternative ideal—of living with conflict gracefully and in full consciousness—gives us a different conception of what 'being at peace with oneself' and 'inner harmony' (Eagle 2013: 908) might come to.

(¹⁹) I argue for this point on other grounds as well in Leite 2016.

(²⁰) There is a parallel issue that can arise regarding psychoanalytic empathy—the analyst's capacity to empathically share in the patient's perspective without endorsing it and while maintaining her own.

(²¹) Of course, this characterization overlooks the vast differences between Frankfurt's and Korsgaard's views in other respects.

(²²) I am grateful to many people over many, many years for assistance related to this chapter. I would especially like to thank Kate Abramson, David Bell, David Bleecker, Matt Boyle, Louise Braddock, Jason Bridges, Ronald Britton, Fred Busch, Morris Eagle, Gary Ebbs, Peter Fonagy, Richard Gipps, Louise Gyler, Edward Harcourt, Jim Hopkins, Leon Kleimberg, Michael Lacewing, Jonathan Lear, Richard Moran, Caroline Polmear, and Mary Target. I apologize to the many additional people who I am surely forgetting. I would also like to thank the audiences at the 2013 Auburn University Conference on Theoretical Rationality, the 2014 University of London/London Institute of Psychoanalysis Joint Conference on Philosophy and Psychoanalysis, the 2015 American Philosophical

Association Pacific Division Meeting, and the members of the Philosophy and Psychoanalysis Study Group of the London Institute of Psychoanalysis.

Adam Leite

Department of Philosophy, Indiana University, Bloomington, USA

Oxford Handbooks Online

Making the Unconscious Conscious 🔓

David H. Finkelstein

The Oxford Handbook of Philosophy and Psychoanalysis
Edited by Richard G.T. Gipps and Michael Lacewing

Subject: Philosophy, Philosophy of Mind   Online Publication Date: Nov 2018
DOI: 10.1093/oxfordhb/9780198789703.013.18

Abstract and Keywords

How should we understand the distinction between conscious states of mind and
unconscious ones? This chapter briefly reviews an answer to this question that the author
has set out and defended in earlier work; it then suggests a new answer—one that
supplements, rather than replaces, the old answer. In spelling out this new answer, the
chapter offers an account of a distinction that is related to, but not identical with, that
between conscious and unconscious states of mind, viz. the distinction between conscious
and unconscious expressions.

Keywords: expression, expressivism, conscious, consciousness, first-person authority, self-knowledge, unconscious

1. A Question

How should we understand the distinction between conscious states of mind and
unconscious ones? This question, which I'll refer to as 'Q' in what follows, could also be
put this way: by virtue of what is someone's anger, fear, anxiety, or desire, for example,
rightly characterized as either conscious or unconscious? In what follows, I'll introduce
some background to, and then briefly review, an answer to Q that I've set out and
defended in earlier work.[1] I'll raise a worry about how this answer could be correct, and
I'll discuss an exchange between Jonathan Lear and Richard Moran that bears on how the
answer I've given to Q might help to elucidate what Lear calls 'the psychoanalytic
meaning of making the unconscious conscious' (Lear 2011: 52).[2] Both the worry and the
exchange between Lear and Moran will lead me in the direction of a new answer to Q—
i.e. toward another way of understanding the difference between conscious and
unconscious states of mind—one that supplements, rather than replaces or corrects, the
old answer. Providing it will require that I present an account of a distinction that is
related to, but not identical with, that between conscious and unconscious states of mind,
viz. the distinction between (what I'll call) conscious and unconscious expressions.

Making the Unconscious Conscious

A good point of departure for theorizing about what it means to characterize someone's state of mind as either 'conscious' or 'unconscious' is a simple, tempting, and ultimately unsatisfactory answer to Q that can be put as follows: 'A person's mental state is conscious if she is aware of it, i.e. if she knows that she is in it. It is unconscious if she does not know that she is in it'. To see why this is not a satisfactory answer to Q, consider the following story:

> Max is a philosophy professor who lives and works in Chicago. Every few months, his mother flies to Chicago from her home in New York to meet with business associates and to visit her only child. Almost every time she does this, Max forgets to pick her up at the airport. Now, as a rule, Max doesn't tend toward this sort of forgetfulness. He never forgets to retrieve his wife or his friends from the airport. Moreover, on those occasions when he forgets to retrieve his mother, he also manages to be away from his mobile phone or to be carrying a phone whose battery has died—so, she is unable to reach him. One day, while talking with a colleague about his mother, it occurs to Max that this behaviour of his might constitute evidence that he is unconsciously angry at her. Indeed, on thinking through some of his other recent behaviour involving his mother, Max becomes convinced by this hypothesis—convinced on the basis of behavioural evidence that he harbours unconscious anger toward her. He says to his colleague, 'I must be unconsciously angry. It's the only way to explain the way I've been acting; I'm angry at my mother'.

Imagine that Max is justified in drawing the conclusion that he does. In such a case, we might describe him as 'knowing that he is unconsciously angry at his mother'. I take it that there is nothing incoherent in such a description. But if this is so, it means that being consciously angry is not the same as knowing that one is angry. At the end of the story, Max knows that he is angry at his mother; he is aware of his anger, but this anger is still unconscious.

Why is it tempting to think that a conscious mental state is just one that its subject knows about or is aware of? I suspect that part of the answer lies in the fact that we use the expression 'conscious of' (nearly) interchangeably with 'aware of'. Right now, as I write these words, I'm aware of—i.e. conscious of—my dog's left front paw, which is resting on my right foot. Of course, no one imagines that it somehow follows from the fact that I'm conscious *of* Kopi's paw that her paw can itself be described as 'conscious'. But the difference between 'conscious *of* x' and 'x's *being* conscious' is easier to get confused about when x is one's own state of mind. In order to get ourselves into a position from which we can think clearly about question Q, we need to distinguish between two ways in which the word 'conscious' is used. At the conclusion of Max's story, he has become aware of—we could say 'conscious of'—his anger toward his mother. But his anger is not conscious. He is conscious *of* his anger, but he is not conscious*ly* angry. (It is only when someone is conscious*ly* angry—and not merely conscious of his own anger—that his anger is said to *be* conscious). Q asks how we should understand the difference between conscious states of mind and unconscious ones (i.e. the difference between someone's

being consciously, e.g. angry and his being unconsciously angry). We'll fail to get this difference into focus if we conflate it with a distinction between things that one is aware of and things that one is unaware of.[3]

2. Wittgenstein's Suggestion

The answer to Q that I've defended in the past may be understood as an attempt to develop a suggestion that Ludwig Wittgenstein makes concerning how to think about psychological self-ascriptions. In his late writings, Wittgenstein often suggests that we should understand such ascriptions as, or as akin to, *expressions*. Thus he writes:

> The statement "I am expecting a bang at any moment" is an *expression* of expectation.
>
> (Wittgenstein 1967: §53)
>
> When someone says "I hope he'll come"—is this a *report* about his state of mind, or an *expression*[4] of his hope?—I can, for example, say it to myself. And surely I am not giving myself a report.
>
> (Wittgenstein 1953: §585)
>
> For even when I myself say "I was a little irritated about him"—how do I know how to apply these words so precisely? Is it really so clear? Well, they are simply an expression.
>
> (Wittgenstein 1992: 70)

Wittgenstein thinks that, at least when we're doing philosophy, we are inclined to imagine that saying, 'I expect an explosion', is more like saying, 'He expects an explosion', than in fact it is. Thus, we assume that just as my ascribing an expectation of an explosion to another person requires that I have some epistemic ground or justification for believing that he expects an explosion, so, too, my ascribing an expectation of an explosion to *myself* requires that I have some epistemic ground or justification for believing that I expect an explosion. The likely result will be a view according to which, even under the best of circumstances (even when there is no question of my being opaque to myself), psychological self-ascriptions should be understood as reports of observations—a view that Wittgenstein takes to be misguided.[5] His suggestion—that we think of such ascriptions as expressions—is meant to provide an alternative way of understanding their grammar. When Wittgenstein calls, e.g. a self-ascription of expectation an 'expression', part of what he means is that it is not the report of an observation. Thus, there is a fundamental grammatical asymmetry between psychological self-ascriptions and ascriptions of psychological states and goings-on to others.

In Wittgenstein's *Philosophical Investigations*, this suggestion first appears in §244, in the context of an exchange between Wittgenstein and his interlocutor concerning how it is

that children manage to learn the names of sensations. During this exchange, we find Wittgenstein saying:

> Here is one possibility: words are connected with the primitive, the natural, expressions of the sensation and used in their place. A child has hurt himself and he cries; and then adults talk to him and teach him exclamations and, later, sentences. They teach the child new pain-behaviour.
>
> (Wittgenstein 1953: §244)

It's no accident that it is in the context of a discussion of language learning that Wittgenstein introduces the idea that some self-ascriptions might be understood in terms of expression. Part of his point could be put as follows (and here I'm reading §244 through the lens of many other passages in which Wittgenstein likens self-ascriptions to expressions). No one imagines that when an infant expresses a pain by crying, wincing, or moaning, he is issuing a report concerning an item that he has discovered via some sort of inwardly directed observation. We can think of a child's gaining the ability to say, 'Ouch!' or, 'That hurts!' or, 'I feel a pain in my foot', as a matter of his learning new ways to do something that, in some sense, he could already do before he could talk, viz. express his pains. And while an expression that takes the form of a prelinguistic cry or wince *is* different in significant ways from an expression that takes the form of a self-ascription[6]—even so, Wittgenstein thinks, there is no good reason to hold that when a child becomes able to produce the latter sort of expression, he suddenly needs to rely on inner observation in order to manage it.

Immediately following the bit of *Investigations* §244 that I quoted in the preceding paragraph, the text continues as follows:

> "So you are saying that the word 'pain' really means crying?"—On the contrary: the verbal expression [*der Wortausdruck*] of pain replaces crying and does not describe it.

The question that Wittgenstein's interlocutor asks in the first of these two sentences is liable to seem oddly confused or unmotivated. After all, Wittgenstein has not suggested (in §244 or elsewhere) that the word 'pain' *means* something different from what we always took it to mean; he's not suggested that it really refers to behaviour. What is moving the interlocutor to ask what he asks here?

Wittgenstein's interlocutor is in the grip of a false dilemma concerning sensations and other mental items. He (the interlocutor) consistently thinks—not only in the *Investigations*, but in other late writings as well—that one must choose between (1) thinking of pains, etc. as things that can be known about only via a kind of inner observation directed at a realm of objects that just one subject can, in principle, have direct access to, and (2) embracing a kind of behaviourism that, in effect, *eliminates* sensations and replaces them with behaviour. In the quoted passage from §244, the interlocutor gathers that Wittgenstein means to be rejecting the idea that we ordinarily

self-ascribe pains on the basis of inner observation, so he, in effect, accuses him of embracing behaviourism. In this instance, the interlocutor's *way* of accusing Wittgenstein of embracing behaviourism is to suggest that Wittgenstein thinks the word 'pain' doesn't after all pick out anything genuinely inner, but instead just means crying.

This isn't the only moment in the *Investigations* when Wittgenstein's interlocutor accuses him of behaviourism. But in §244, Wittgenstein replies to the accusation in a way that is liable to sound off-key. He doesn't say merely that an utterance such as 'I'm in pain' should be understood as a verbal expression of pain and not as a description of crying (or of any other behaviour). He says that the verbal expression of pain *replaces* ('ersetzt') crying. And perhaps this doesn't seem quite right. After all, it's not as if children *stop* expressing pains by crying when they learn how to express them by talking about them.

Maybe Wittgenstein ought to have used a different word at the end of §244. Still, I think it's important that there is, after all, a *kind* of replacement that occurs when children learn to express their pains by self-ascribing them. And I think Wittgenstein *might* have had this in mind when he imagined the exchange in the way that he did. In any case, *I* want to say that there is an important sense in which the *kind* of crying a child did as a prelinguistic infant really is no longer a part of her life after she learns to talk about her own sensations, emotions, and attitudes. The kind of crying she did as an infant is, at that point, replaced by another kind of crying. After she learns to talk, even though her expressions-of-pain-via-crying sometimes look and sound the same as they did before, they now could be said to have a different *form*. I'm going to have to leave this suggestion dark for a little while—until after I've discussed the possibility of an expression's being conscious. I'll come back to it in §7.

3. An Answer to Q

In my own writing, I've tried to take what Wittgenstein says concerning the grammar of our talk about our own mental states and use it to make sense of the distinction between two kinds of mental state. The guiding idea could be put this way: once a human being has acquired language, her mental states are such that she is, ordinarily, able to express them via the sort of speech act that Wittgenstein talks about; she can express her sadness, anger, or desire by ascribing it to herself. Insofar as she is able to do this, her sadness, anger, or desire is said to be *conscious*. But sometimes, even after a human being has acquired language (and with it, the capacity to say what she is thinking or feeling), she is unable to express, e.g. some particular fear of hers by self-ascribing it. A kind of expressive ability that is characteristic of normal psychological states and events as they figure in the life of a linguistic animal is, in such a case, absent or blocked. The word that we use to characterize this sort of absence is 'unconscious' or 'unconsciously'; we say that someone's fear is unconscious or that she is unconsciously afraid.

What would it take for Max's unconscious anger toward his mother to *become* conscious? It would not be sufficient for Max to become aware on the basis of behavioural evidence that he was angry. For Max's anger to become conscious, he would need to acquire, or to

regain, an ability to express it merely by self-ascribing it. This expressive ability is what he still lacks at the conclusion of the story I told about him in §1. At the end of that story, Max says to a colleague, 'It's the only way to explain the way I've been acting; I'm angry at my mother'. In so saying, Max expresses a belief (or perhaps knowledge) that he has about himself. But he does not express his anger at his mother. He may express that anger by stranding his mother at the airport, but he is still unable to express it by self-ascribing it. Hence, it is still unconscious.[7] This is to say, I've answered Q by claiming that what distinguishes conscious states of mind and unconscious ones is the presence or absence of a particular expressive ability.

In my *Expression and the Inner*, I put the point as follows:

> *Someone's mental state is conscious if he has an ability to express it merely by self-ascribing it. If he lacks such an ability with respect to one of his mental states, it is unconscious.*

> (Finkelstein 2003: 120)

What work is the word 'merely' doing in this encapsulation of an answer to Q? Why not leave out 'merely' and say, more simply, that a person's mental state is conscious if and only if he can express it in a self-ascription? The trouble with the simpler formulation is that it is open to counter-examples of the following sort. Imagine that Max occasionally expresses his anger by speaking in a peculiar, clipped tone of voice. *While speaking in this tone of voice*, he says: 'The only way to explain my odd behaviour is to posit that I am unconsciously angry at my mother. So, I'm angry at her'. Through his tone of voice, Max expresses his anger in a self-ascription of it. Even so, his anger is unconscious. In *Expression and the Inner*, I wrote:

> When I say that someone's state of mind is conscious if he has an ability to express it *merely* by ascribing it to himself, I mean this: the sort of ability at issue is one that enables a person to express his state of mind in a self-ascription of it, where what matters [to the self-ascription's expressing what it does] isn't his tone of voice (or whether he is tapping his foot, or what he is wearing, or to whom he happens to be speaking), but simply the fact that he is giving voice to his sincere judgment about his own state of mind.[8] That I might manage to express my unconscious anger in a self-ascription of it *via a clipped tone of voice* doesn't show that I have the relevant sort of expressive ability. When I am consciously angry, I can say in a neutral tone of voice, "I'm angry," and thereby express my anger.

> (Finkelstein 2003: 120)

Was that an adequate response to the tone-of-voice case? In one sense, I think it was. It showed that the case constitutes a merely apparent, not a genuine, counter-example to the view I was defending in the book (and that I still take to be correct). On the other hand, it seems to me now that what I wrote there (and elsewhere) is liable to leave a reader with a nagging question that might be put as follows: 'Why should it *matter* if a

Subscriber: Goldsmiths College, University of London; date: 08 March 2019

person is able to express his anger merely by self-ascribing it—as opposed to being able to express it in a self-ascription that comes in a clipped tone of voice? How *could* the difference between conscious and unconscious anger come down to something so apparently unimportant as that?' I take the challenge posed by this question seriously. Indeed, part of what has moved me to pursue a second answer to Q is the aim of coming to better understand how, given this sort of challenge, my first answer could be correct. How is it that gaining, or regaining, an ability to express some emotion or attitude merely by self-ascribing it can make a significant difference in a person's life? I'm not yet in a position to address this question, but I'll return to it after I've given a second answer to Q, in the final section of this chapter.

4. Non-Linguistic Animals

Given the answer to Q that I've just been reviewing, we should not expect the distinction between conscious and unconscious psychological states to, as it were, *get a grip* when we are thinking about non-linguistic animals. After all, it is only when one is considering the psychology of a linguistic creature that it makes sense to distinguish between states that can be expressed in self-ascriptions and states that cannot. And, as a matter fact, we *don't* characterize the psychological states of non-linguistic creatures as either conscious or unconscious. Even those of us who are prone to attribute complex thoughts and feelings to our pets do not say things like: 'My dog gets upset when I pack my suitcase because she has an unconscious fear of being abandoned', or, 'Over time, my cat's distaste for the people who live next door has gone from being unconscious to being conscious'. Nor do we say such things about prelinguistic children. To borrow a phrase that Wittgenstein uses in a related context,[9] when a child learns to attribute mental states to herself, a 'new joint' is added to the language-game. We could say that when children learn to self-ascribe their psychological states and events, the *form* of those states and events changes. From then on, they are either conscious, unconscious, or somewhere in between conscious and unconscious.[10] To the extent that some desire of mine is such that I am unable to express it by self-ascribing it, it could be said to suffer from a formal defect or imperfection. But it is no imperfection in Kopi's desire, e.g. for water that she is unable to express it by self-ascribing it. Kopi's desires are, unlike mine, neither conscious nor unconscious.[11] This is to say: it makes no sense to describe a desire of hers as either 'conscious' or 'unconscious'. In the shift from brute to fully realized human desire, a new joint is added to the language-game; the psychological is transformed.

5. 'What is gained ... by replacing some words for some tears?'

One of the goals that Jonathan Lear sets for himself in his fascinating *A Case for Irony* is to elucidate 'the psychoanalytic meaning of making the unconscious conscious' (Lear 2011: 52). In pursuing this end, he draws on the strand in my work that I've been reviewing and discussing in this chapter. Lear contrasts a case in which someone

concludes on the basis of self-observation that he must be angry, even though he isn't 'consciously feeling angry', with a very different case that he describes as follows:

> [I]n the midst of a boiling rage, I say "I'm *furious* with you!" In this case, the anger itself is present in the verbal self-ascription of anger that is directed at you. In this case, the self-ascription *of* anger is itself an *angry* expression. Unlike the former case, I do not have to observe myself to know that I am angry. I just *am angry*; and the form my anger at you takes on this occasion is the angry verbal self-ascription directed at you. In this case, the utterance "I am *furious* with you!" may replace other forms of angry expression. ... This is a case, in Finkelstein's terms, in which I am not only conscious *of* my anger; I am *consciously angry*.
>
> (Lear 2011: 52–3)

Lear goes on to consider the question of how it might be therapeutic, rather than merely disruptive, to find some way of, as it were, taking unconscious anger—anger that has been 'held out of consciousness not simply because it is painful, but because it violates one's sense of who one is' (Lear 2011: 54)—and making it conscious. In this context, he says, 'The question thus becomes whether there could be a process—not *too* disruptive— by which the verbal expression of anger *replaces* (in the Wittgenstein-Finkelstein sense) its unconscious manifestations' (Lear 2011: 54).

A Case for Irony comprises a pair of lectures by Lear followed by responses by several other philosophers. In one of these responses, Richard Moran writes:

> The idea of "replacing" one mode of expression for another is not perfectly clear to me, particularly in the therapeutic context into which Lear is importing the idea. On a basic level, we might ask: If the two modes of expression are really doing the same work, then what is the *point* of "replacing" one with the other? If we take the notion of replacement literally, then what is gained, for instance, by replacing some words for some tears? Does that mean that the tears are now unnecessary, having been replaced by something else?
>
> (Moran 2011: 113)

It might be tempting to try to answer Moran's question by drawing a distinction between the kind of expressive 'work' that gets done via a self-ascription of sadness and the kind that gets done via crying. But I believe that this would be to draw a line in the wrong place. At least insofar as we're interested in psychoanalysis, there *is* an important line for us to draw between expressions that are, in a very particular way, defective or imperfect and ones that are not. But, as will become apparent, crying can lie on *either* side of this line. Moran's question is useful, I think, insofar as it helps us to see that we need to distinguish cases in which someone's (e.g.) crying constitutes (what can be called) a conscious expression of sadness (or joy or whatever) from ones in which someone unconsciously expresses her (e.g.) sadness via crying. Moran asks what the point is of replacing some tears with some words—that is, of replacing an expression that takes the

form of crying with one that takes the form of self-ascription. My own view is that there need not be *any* point in replacing crying with self-ascription. Tears are just fine as long as they constitute *conscious* expression. Psychoanalysis does not aim to replace tears with words. It aims to replace unconscious expressions with conscious ones.

6. The Distinction between Conscious and Unconscious Expressions

What does an unconscious expression look like? We already saw an example of one in §1, but I'd like to consider a somewhat different example now:

> One afternoon, Max and his wife, Sarah, have a quarrel while assembling hors d'oeuvres for a dinner party that they'll be hosting that evening. Sarah's new boss will be one of their guests, and partly for this reason, she is worried about the party's going well. Although the quarrel subsides after a few minutes, Max continues to dwell on it. Just before guests are scheduled to arrive, he reminds himself that Sarah has been anxious about the party's going smoothly, and he tells himself that it would be petty of him to express the anger that he's still feeling toward her in front of their guests. He resolves to be the perfect co-host, to be cheerful, charming, and useful. In spite of—and perhaps partly because of—this resolution, Max repeatedly expresses his anger during the party, but he remains unaware of doing so. He responds to Sarah's sharp, witty, and critical assessment of a recent movie by suggesting that she simply failed to appreciate what was innovative about it; he makes disparaging jokes about food that Sarah prepared; he tells a story about her brother that she finds embarrassing; and so on. It does not occur to Max that these remarks of his might be connected to that afternoon's quarrel, and if someone were to suggest that there *was* a connection, he would deny it. Nonetheless, it is apparent to all but the drunkest of their friends that Max is annoyed with Sarah. For all that, Max performs the bulk of the serving and most of the straightening up. Indeed, at one point during the evening, he privately congratulates himself on managing to 'put on a good face' for Sarah's sake, in spite of his angry, hurt feelings.

Here, Max unconsciously expresses his anger toward Sarah in passive-aggressive remarks. A noteworthy feature of this particular example of unconscious expression—a feature that is not present in every example—is that here, a person unconsciously expresses a state of mind that is, itself, conscious. (As I've asked you to imagine him, Max is consciously angry at Sarah). This example stands in contrast to one in which someone unconsciously expresses an unconscious state of mind. The story that I told in §1—in which Max expresses his unconscious anger by repeatedly stranding his mother at the airport—constitutes one such example.

Thus far, I've recounted two stories in which Max *unconsciously* expresses anger. Here's one in which he *consciously* expresses pleasure:

> Early one evening, Max and Sarah are feeling both hungry and tired, so rather than cooking dinner, they decide to go to a nearby restaurant. Five minutes after they order, Sarah tells a joke, and after a brief pause, Max's face lights up in a smile. Sarah says, 'Are you smiling at my joke?' Max replies, 'No. I thought your joke *was* funny, but I'm smiling because I see the waiter with our food, and I'm really pleased that it's coming out so quickly'.

In this example, Max is able to speak to the question of what his smile expresses in, we could say, a first-personal way. His smile is a conscious expression of his pleasure at the imminent arrival of dinner.

How is conscious expression possible? How is it that Max is able to just say what his smile expresses? How does he know that he's *not* smiling at Sarah's joke? I want to approach these questions by first considering what seems to me an easier sort of case, one in which someone speaks with a kind of first-person authority—not about what psychological state a bit of his behaviour expresses, but rather—about what some of his own words mean:

> One afternoon, Max and Sarah attend a philosophy department social event. A graduate student says to Max: 'I just met Sarah. Is she a philosopher too?' Max replies: 'She *is* a philosopher, by which I mean: she is moved by, and thinks hard about, philosophical questions. But she's not someone who makes her living by doing philosophy'.

Here, Max first says, 'Sarah *is* a philosopher', and then he provides a gloss on what he meant by these words. Imagine that someone were to ask: 'How did Max find out what he meant by the words, "Sarah *is* a philosopher"? Did he first hear himself utter these words and then work out a plausible interpretation of them?' That's the wrong way to think about what's going on in the story. Indeed, the question, 'How did Max find out what he meant when he said, "Sarah *is* a philosopher"?' is itself confused. Max didn't *find out* what the first half of his statement meant and then, in the second half of the statement, report what he had learned. We might want to call the second part of Max's statement an interpretation of, or gloss on, the first part. But it's not the sort of gloss or interpretation that another person might provide. The second part of Max's statement is as much a *part* of his assessment of Sarah *qua* philosopher as the first part is. We should think of the two parts of Max's statement as constituting a kind of unity, a single act. I'll describe them as *co-expressing* his assessment of Sarah *qua* philosopher. If a listener wants to understand this assessment, she needs to consider the two parts of Max's statement (the part in which he says, 'Sarah *is* a philosopher', and the part in which he glosses what he meant by these words) as making sense together, as two pieces of a single, extended expression of his thought.

Now consider again the story I told about Max and Sarah at the restaurant. Recall that when Sarah asks Max whether he's smiling at her joke, he replies, 'I'm smiling because I see the waiter with our food, and I'm really pleased that it's coming out so quickly'. Max's smile, I said, is a conscious expression of his pleasure at the imminent arrival of dinner. I

want to suggest that we understand Max's gloss on what his smile expresses in the same way that we understood Max's gloss on what he meant by the words, 'Sarah *is* a philosopher'. When Max says what his smile expresses, he is not offering the sort of interpretation of his behaviour that another person might offer. Rather, his smile and his gloss on its significance are two parts of a unified expressive act; together, the smile and the gloss co-express his pleasure at seeing dinner on its way. What I'll call a 'co-expressive gloss' on the psychological significance of, e.g. a smile is not an interpretation of the smile based on observation. It is, rather, a kind of extension of the smile into a linguistic register.

***

At this point, I can say what it is that distinguishes conscious expressions from unconscious ones. Someone's expression of a mental state is conscious if she is able to gloss it—i.e. gloss its psychological significance—co-expressively. Someone's expression of a mental state is unconscious if she is unable to gloss it co-expressively. What it means to say that Max's passive-aggressive dinner-party remarks are unconscious expressions of his anger at Sarah is this: they are expressions of anger at Sarah that Max is unable to gloss co-expressively.

It's not impossible that someone should unconsciously express her frustration by slamming a door and, immediately thereafter, accurately interpret the significance of this behaviour in the way that a psychologically astute external observer might. But in such a case, the door-slam and the gloss on it would be separate acts. When someone consciously expresses her frustration by slamming a door, she is able to provide a gloss on the slam, where such a gloss would, itself, be a continuation—a kind of extension or stretching out—of her expressive act.

7. Replacement

Earlier in this chapter (§2), I discussed §244 of *Philosophical Investigations*, in which Wittgenstein introduces the idea that when a child learns to talk about his pains, he thereby learns a new way to express them. I noted that §244 concludes in a way that's liable to sound off-key, with Wittgenstein saying to his interlocutor, '[T]he verbal expression of pain replaces crying'. I went on to make a suggestion that I was not yet in a position to spell out, viz. that even if 'replaces' ('ersetzt') seems the wrong word for Wittgenstein to have used in this context, there *is* a sense in which the kind of crying that children do as prelinguistic infants is—when they learn to talk—replaced by another kind of crying. I want to return to that suggestion and try to explain it in light of what I just said about the distinction between conscious and unconscious expressions along with something I said a little while ago, in §4.

In §4, I noted that the distinction between, e.g. anger that can be expressed via self-ascription and anger that cannot be so expressed does not, as it were, get a grip when one is considering the mind of a non-linguistic or prelinguistic animal. I went on to say

that when a child gains the ability to self-ascribe his attitudes and emotions, their form changes. Until then, it doesn't make sense to describe them as conscious or as unconscious. From then on, they are either conscious or unconscious or in between conscious and unconscious. Now, something similar happens when a child gains the ability to co-expressively gloss the psychological significance of his expressions: the form of his *expression* changes. Until then, it doesn't make sense to describe his expressions (of sadness, pleasure, love, desire, anger, etc.) as either conscious or as unconscious. From then on, his expressions are either conscious or unconscious or in between. When Kopi expresses a desire to play with me by dropping a Frisbee at my feet, it is no formal imperfection in her expression that she is unable to extend it into a gloss on itself. But if Sarah discovers Max curled up on their bed, sobbing, and if—when she asks him why he is crying—he is completely unable to say, then his crying could be characterized as imperfectly instantiating the form of human expression. Thus, when children learn to talk, one kind of crying does, after all, replace another.

Perhaps this a good place to point out that in this chapter, I am discussing four distinct kinds of *replacement*: (i) the replacement of prelinguistic desires, fears, etc. with the sort of desires, fears, etc. that are characteristic of linguistic creatures, (ii) the replacement of prelinguistic expressions of desire, fear, etc. with the sort of expressions of desire, fear, etc. that are characteristic of linguistic creatures, (iii) the replacement of unconscious desires, fears, etc. with conscious ones, and (iv) the replacement of unconscious expressions of desire, fear, etc. with conscious ones. A case of type (i) or (ii) does not constitute an example of 'making the unconscious conscious'.[12] Cases of types (iii) and (iv) *are* examples of 'making the unconscious conscious'. A central aim of this chapter is to elucidate the logical relations between cases of types (iii) and (iv)—between, that is, two different ways in which the unconscious may be made conscious.

8. Another Answer to Q

In the story that I told about Max and Sarah's dinner party, Max expresses his anger at Sarah via passive-aggressive remarks. As I noted in §6, that story illustrates that it is possible for someone to unconsciously express his conscious anger. Indeed, there is nothing particularly extraordinary about a person's unconsciously expressing either an unconscious emotion or a conscious one. By contrast, we don't find cases that we'd be inclined to describe as ones in which someone *consciously* expresses an *unconscious* state of mind. This is what we should expect, given what I have claimed about the distinction between conscious and unconscious expressions. In order for Max to consciously express, e.g. his anger toward his mother, he would need to produce an expression of anger whose psychological significance he could co-expressively gloss. But now: consider a case in which Max expresses his anger toward his mother by not answering the phone when he sees that she's calling. If Max *could* co-expressively gloss this expression—if he could say, 'I'm not answering the phone because I'm angry at my mother', and, thereby, extend his non-verbal expression into a gloss—his anger would be conscious.

Summing up a bit: we've seen that conscious states of mind may be expressed either consciously or unconsciously. But unconscious states of mind can be expressed only unconsciously. There is, as it were, no grammatical or logical space for an unconscious emotion or attitude to be consciously expressed. A bit of reflection on this point yields the new answer to Q that I promised at the start of this chapter. In §3, I quoted the following statement from *Expression and the Inner*:

> *Someone's mental state is conscious if he has an ability to express it merely by self-ascribing it. If he lacks such an ability with respect to one of his mental states, it is unconscious.*

> (Finkelstein 2003: 120)

Let's call that first answer to Q 'A1'. What follows is a second answer; call it 'A2'.

> *Someone's mental state is conscious if she has an ability to consciously express it. If she lacks this ability with respect to one of her mental states, it is unconscious.*

A2 might sound as if it involves an uncomfortably tight circle, so I should point out that although I am here, as it were, defining a sort of 'consciousness' in terms of a sort of 'consciousness', the latter sort of consciousness has been further defined in terms of co-expressive glossing.[13]

Although A1 and A2 may be thought of as providing two different criteria for a mental state's counting as conscious, if they both are true—as I believe they are—then any mental state that meets either one of these criteria will meet the other one as well. This is to say: (i) if someone is able to express his, e.g. anger merely by self-ascribing it, then he will be able to consciously express it. And (ii) if someone is able to consciously express, e.g. his anger, then he will be able to express it merely by self-ascribing it.[14] In what remains of this chapter, I'll be particularly interested in a point that is implied by (ii), viz. that if someone is *unable* to express his, e.g. anger merely by self-ascribing it, then he will be *unable* to consciously express it.

***

Given the answer to Q that I summarized in §3 and that is encapsulated in A1, making an unconscious emotion or attitude conscious is a matter of acquiring or regaining an ability to express it merely by self-ascribing it. It is such an ability that, I said, Max still lacks even after he's concluded on the basis of behavioural evidence that he is angry at his mother. At the end of §3, I introduced, and then left hanging, a question concerning how this could be the right thing to say about what it takes for an unconscious state of mind to become conscious. Someone might put this question to me as follows: 'As you have characterized him, Max is able to express his unconscious anger at his mother by stranding her at the airport, by misplacing his phone when he knows that she is likely to call, and, presumably, by doing a wide variety of other things that are liable to annoy her. Given the long list of ways in which he can *already* express this anger, why should his

becoming able to express it in *one more way*—by self-ascribing it—be significant? How could the difference between unconscious and conscious anger come down to *that*?'

I am now in a better position than I was in §3 to address the challenge presented by this question. A couple of paragraphs back, I noted that putting A1 and A2 together yields a point that I put as follows: 'if someone is *unable* to express his, e.g. anger merely by self-ascribing it, then he will be *unable* to consciously express it'. I suggest that this point ought to help us understand why it is that the ability to express one's own state of mind in a self-ascription of it should *not* be considered just one expressive ability on a long list of such abilities. Until Max is able to express his anger at his mother merely by self-ascribing it, there is no way for him to consciously express it at all—either by telling his mother that he is angry at her or by refusing to let her stay with him and Sarah when she visits or by choosing not to answer the phone when he sees that she is calling.[15] If Max were to gain the ability to express his anger at his mother merely by self-ascribing it, this would, as it were, open up the possibility of his consciously expressing it in these and myriad other ways. If only for this reason, such a shift would be significant; it would matter to Max's life.

For all that, acquiring an ability to express his anger toward his mother in a self-ascription of it would not guarantee that Max's subsequent expressions of this anger would be conscious rather than unconscious. As we've seen, a conscious state of mind may be expressed *either* consciously or unconsciously. At the end of §5, I said that psychoanalysis aims—not to replace tears with words, but—to replace unconscious expressions with conscious ones. It is perhaps disappointing that acquiring (or regaining) an ability to express an emotion or attitude in a self-ascription of it does not ensure that it won't continue to be expressed unconsciously. Still, the acquisition of such an ability is a necessary step toward replacing unconscious expressions with conscious ones.[16]

References

Finkelstein, D. (1999). 'On the distinction between conscious and unconscious states of mind'. *American Philosophical Quarterly 36* (2) (April): 79–100.

Finkelstein, D. (2003). *Expression and the Inner*. Cambridge, MA: Harvard University Press.

Finkelstein, D. (2010). 'Expression and avowal'. In Kelly Jolley (ed.), *Wittgenstein: Key Concepts* (pp. 185–198). Durham: Acumen Press.

Kluger, J. (2017). *The Animal Mind*. New York: Time Inc. Books.

Lear, J. (2011). *A Case for Irony*. Cambridge, MA: Harvard University Press.

Moran, R. (2011). 'Psychoanalysis and the limits of reflection'. In Jonathan Lear, *A Case for Irony* (pp. 103–114). Cambridge, MA: Harvard University Press.

Wittgenstein, L. (1953). *Philosophical Investigations*, trans. G. E. M. Anscombe. Oxford: Basil Blackwell.

Wittgenstein, L. (1967). *Zettel*, trans. G. E. M. Anscombe, eds. G. E. M. Anscombe and G. H. von Wright. Oxford: Basil Blackwell.

Wittgenstein, L. (1992). *Last Writings on the Philosophy of Psychology Volume 2*, trans. C. G. Luckhardt and M. A. E. Aue, eds. G. H. von Wright and H. Nyman. Oxford: Blackwell.

Notes:

(1.) See, e.g. Finkelstein 1999 and Finkelstein 2003.

(2.) Although Lear is interested in how my answer to Q might bear on our understanding of psychoanalysis, it's worth mentioning that neither Q nor the answer to it that I've given is *about* psychoanalysis as such. (Freud was not the first writer to recognize that we are sometimes moved by unconscious emotions, preferences, and desires. I take it that the title character in Jane Austen's *Emma*, who realizes at the end of the novel that she's been in love with Mr. Knightly for quite a while, presents at least as good an example of someone whose behaviour is motivated by unconscious states of mind as does, say, Freud's Rat Man).

(3.) I believe that many theorists of 'consciousness' fall into confusion by failing to distinguish between 'conscious of x' and 'x's being conscious'. For a discussion of this point, see Finkelstein 2003, ch. 1.

(4.) In her translation of *Philosophical Investigations*, G. E. M. Anscombe sometimes renders 'äußerung' as 'expression' and sometimes as 'manifestation'. Here, I've changed her 'manifestation' to 'expression'.

(5.) It would take me beyond the scope of what I aim to do in this chapter to explain why Wittgenstein thinks this is misguided. Stepping back from Wittgenstein, I would suggest that any such position will suffer from the conflation that I mentioned in n. 3.

(6.) One difference is that winces, moans, and cries (regardless of whether or not they are produced by a human being who can already talk) are not truth-apt, while self-ascriptions are. Wittgenstein has often been read as wanting to deny this difference and claim that psychological self-ascriptions are neither true nor false. This reading is, I think, both uncharitable and unsupported by the relevant texts. (For a discussion of this point, see Finkelstein 2010).

(7.) 'What if someone has a phobia that prevents him from ever speaking about being angry? Couldn't he nevertheless be consciously angry?' As I understand the concept of expression, not every expression must be a *public* manifestation. (It seems to me that on more than one occasion, I have: (i) tried to tiptoe past a sleeping person, (ii) stubbed my toe, and, finally, (iii) expressed the resultant pain in my toe by *thinking*, 'Ow! That really, really hurts!') It's consistent with my understanding of expression to allow that we

Subscriber: Goldsmiths College, University of London; date: 08 March 2019

sometimes express our states of mind by self-ascribing them in thought, rather than out loud. So yes; someone who has a phobia about *saying* that he's angry can, on my view, nevertheless be consciously angry.

(8.) Imagine someone's replying to what I said in this sentence as follows: 'If a person expresses his, e.g. anger by self-ascribing it, then what his self-ascription "gives voice to" is just that anger and *not*, as you suggest here, a judgment about it'. In order to bracket this concern, my former self might have rewritten this sentence so that it ended as follows: 'but simply the fact that he is sincerely addressing a question about his own state of mind'.

(9.) Wittgenstein 1967: §§422–6.

(10.) I argue that we should think of conscious and unconscious states of mind as lying at two ends of a continuum in Finkelstein 2003: §5.5.

(11.) This isn't to deny that Kopi might be very conscious *of*, e.g. a piece of cheese that's sitting on a plate on a table just above her head. Nor is it to deny that she might bang her head—perhaps in pursuit of that cheese—and momentarily *lose consciousness*. (In a special edition of *Time* called *The Animal Mind*, Steven Pinker is quoted as saying, 'It would be perverse to deny consciousness to mammals' (Kluger 2017). When I say that Kopi's states of mind are neither conscious nor unconscious, I don't take myself to be disagreeing with Pinker's remark).

(12.) I've characterized cases of types (i) and (ii) as involving changes in form—*from* states of mind and expressions that cannot be described as either conscious or unconscious *to* states of mind and expressions that can be so described. Thus, cases of these types are examples—not of 'making the unconscious conscious', but rather—of 'making what is neither conscious nor unconscious either conscious or unconscious'.

(13.) I am not claiming that there won't turn out to be *some* sort of circularity involved in one or both of my characterizations of the distinction between conscious and unconscious states of mind. My aim is not to provide a reductive account of consciousness. I don't mind circles, as long as they aren't so small as to be uninteresting.

(14.) If it still seems unclear that (i) and (ii) both obtain, what follows might make it seem less so. Let's begin with (ii). Imagine that Max slams his office door and that this is (for him, in this situation) a conscious expression of anger that he's feeling toward a colleague named Jeff. What it means for this to be a *conscious* expression of anger is that he can co-expressively gloss its psychological significance; he can respond to a question like, 'Why did you slam the door?' with an answer like, 'I slammed it because I'm angry at Jeff'—and he can do this in such a way that the gloss is a kind of extension of (rather than an observation report about) the initial expressive act (the door-slam). Now, if Max can do that, then he is (*obviously*, I want to say) able to express his anger merely by self-ascribing it, i.e. merely by saying, 'I'm angry at Jeff'. So, if Max's anger counts as conscious according to A2, it will also count as conscious according to A1. Thus, (ii) would

seem to be true. Now let's consider (i). Imagine that Max is able to express his anger at Jeff merely by saying, 'I'm angry at Jeff'. Such a self-ascription would *itself* be a conscious expression. This is to say: Max would have the ability to co-expressively gloss *its* psychological significance. Suppose that Max expresses his anger at Jeff by saying to Sarah, 'I'm angry at Jeff'. Sarah thinks (for reasons with which we needn't concern ourselves) that in so saying, Max might be trying to make some sort of joke. She replies, 'Are you joking? I'm not getting what you mean'. Here, Max would be in a position to reply with a co-expressive gloss on his initial expression, as follows: 'When I said, "I'm angry at Jeff", I wasn't making a joke; I was actually expressing anger that I'm feeling'. So, if Max's anger meets the criterion for consciousness in A1, it will also meet the criterion in A2. Thus, (i) would also seem to be true.

(15.) While his anger toward his mother remains unconscious, Max might unconsciously express it by not answering the phone when he sees that she is calling. But until he can express that anger merely by self-ascribing it, there is no possibility of his *consciously* expressing it by not answering the phone when he sees that she is calling.

(16.) Versions of this chapter were presented at the University of Murcia (Workshop on Contemporary Rationalist and Expressivist Approaches to Self-knowledge, October 2016) and at the University of Chicago (2016 Fall Colloquium, Department of Philosophy, November 2016). I'm grateful to members of both audiences for their questions and comments, especially Matthew Boyle, Jonathan Lear, Ángel García Rodríguez, Ryan Simonelli, and Anubav Vasudevan. In addition, I'm indebted to Agnes Callard, Matthias Haase, and Malte Willer for reading and commenting on drafts.

David H. Finkelstein

Department of Philosophy, The University of Chicago

Oxford Handbooks Online

Abstract and Keywords

This introduction provides an overview of the five chapters in this section, which explores some of the central issues in psychoanalysis as it relates to phenomenology and science. One such issue concerns the scientific status of psychoanalysis (natural or social? bona fide or tendentious?), and more specifically which between the methods of psychoanalysis and its real-life practice may be considered scientific. One of the chapters examines psychoanalysis as a scientific theory of mind, arguing that psychoanalysis fails to test its theories. Another chapter suggests that many of the central tenets of psychoanalytic theory are evidentially supported by recent developments in empirical neuropsychology. Also discussed are the debate between those who view psychoanalysis as science and those who insist that it rather offers a hermeneutic, how psychoanalysis provides a phenomenology in its articulations of unconscious life, and alternative phenomenological schemes for framing the dynamic unconscious.

Keywords: psychoanalysis, phenomenology, science, scientific theory of mind, empirical neuropsychology, unconscious, psychoanalytic theory

Michael Lacewing and Richard G. T. Gipps

PHILOSOPHICAL disputation concerning psychoanalysis often begins—and unfortunately sometimes ends—with the question of whether psychoanalysis counts as scientific psychology. The claim that it does not is often taken to imply 'too bad for psychoanalysis', rather than, say, 'too bad for scientific psychology', or 'so what else is it then?' But the debate, if it is to be philosophically respectable, cannot simply take for granted some or other unexamined conception of what being a science amounts to. And it will only generate confusion, rather than pertinent critique, if it imports standards inappropriate to its object. What it needs are explicit, workable, and apt conceptions of scientificity. What we need to know when engaging with the debate is, for example, whether we are here talking narrowly about natural science or much more permissively about

Wissenschaft—i.e. a systematic body of knowledge—or something in between? If by 'science' we mean *Wissenschaft*, then in that case anthropology, natural history, art history, and psychoanalysis shall—as and when they accrue knowledge rather than indulge speculation—all be counted as sciences. Freud certainly used the term Wissenschaft to describe his endeavours, but he also tended to think that 'the intellect and the mind are objects for scientific research in exactly the same way as any non-human things' (1933: 159).

But what does the 'same way' mean in the quotation from Freud? Natural history differs considerably from, say, chemistry in its use of merely observational rather than experimental methods to accrue knowledge. Freud's suggestion assimilates viable understanding of matters psychological to the model of natural sciences; it leaves out of consideration, for example, the possibility that psychoanalysis may better be considered a social science—a science whose objects and methods are importantly unlike those populating, and used to interrogate, non-human nature. We should also consider that even parents with the best intentions and the most delightful offspring can sometimes impose on their children in such a way as risks thwarting rather than nurturing their flourishing. This, at least, is the force of Habermas's (1987) contention that, in relation to his own intellectual progeny, Freud laboured under 'a scientistic self-misunderstanding'.

(p. 350) The question of the scientific status of psychoanalysis (natural or social? bona fide or tendentious?) also requires specification as to the relevant aspect of its object. Are we talking about the metapsychology (e.g. about putative underlying psychological structures such as the id and superego), the aetiological theories (e.g. as concern the developmental origins of various psychopathologies), the clinical theory (e.g. relating particular difficulties to particular current defensive organizations), or the successfulness of the clinical practice (e.g. as evaluated in outcome studies)? And are we talking about whether the methods of psychoanalysis could, in theory, count as scientific or about whether the real-life practice of psychoanalysts does so? And if we are interested in testing the theory, then what kind of procedure is here to count as 'testing'? And if we are talking about the clinical theory, then which aspect of which one? Psychoanalysis has after all developed a large number of theories of various phenomena from which a variety of hypotheses may be drawn for testing, albeit all organized around centrally recurring themes of dynamically unconscious motivation, defence mechanisms, symbolization, and transference.

This last point raises a deeper question about how to understand psychoanalysis. When we are talking about just such fundamentally organizing conceptions as unconscious motivation, etc. are we describing the a posteriori hypotheses of a fallible science, or rather the a priori organizing conceptualizations which belong to something we could call the 'philosophy' of psychoanalysis? (Compare: do concepts such as 'force' and 'space' function to define physicists' very field of enquiry, or do they function by referring, hopefully successfully, to phenomena that show up within that field?) Perhaps that distinction is itself rather too blunt to be helpful, and we should do better to consider the a priori and the a posteriori as separated by degrees rather than by an immutable chasm.

When we meet with an effective piece of clinical psychoanalytical theory, is its effectiveness to be understood in terms of its ability to bring to life or make visible something which beforehand too often went unnoticed—thus enriching the conceptual frameworks with which we articulate human nature, or is it proposing a new causal explanation for something which was perfectly visible all along? With regards to this issue of visibility, a question not often enough asked, which parallels that of the scientific prowess of psychoanalysis, concerns its phenomenological credentials: what wattage of illumination of human nature has really been achieved by psychoanalysis compared with other sciences and humanities?

In this part, the chapter by Eagle focuses on psychoanalysis as a scientific theory of mind. Taking testability as a measure of scientific standing, he takes psychoanalysis to task for its failure to test its theories. While Freudian theories can be used to generate testable hypotheses, Kleinian and post-Kleinian theories are, he suggests, not so readily testable. While they may play a variety of valuable functions inside and outside the clinic, contributing to a scientific theory of the mind is not amongst them. Furthermore, setting aside the theories, the practice of psychoanalysts does not encourage a scientific approach—an objection pressed most frequently by Frank Cioffi (1999). Particular difficulties noted by Eagle include a failure to address confirmation bias, a dearth of decent published case studies, aetiological theories developed merely on the basis of clinical (p. 351) data, concepts being tacitly redefined to preserve the otherwise implausible theories they articulate, and a culture of hostility to research, guild mentality, and gurus.

While Eagle looks at the extant scientific habits of psychoanalysts, Hopkins—a philosopher—aims to demonstrate the scientific significance of psychoanalysis by newly integrating it with attachment theory and neuroscience, arguing that many of the central tenets of psychoanalytic theory are evidentially supported by recent developments in empirical neuropsychology. Hopkins stresses the consiliences between psychoanalytic understandings of the task of assimilating emotionally overwhelming experience (e.g. in dreaming) to maintain reality contact, the development of symptoms when that breaks down and primary process modes of cognition take over, the significance of our attachment histories for developing the neurological structures required for such emotion processing, Freud's early neurological thought on minimizing 'free energy', and the theoretical neuroscientist Karl Friston's more recent conception of the brain as tasked with free energy minimization by producing such 'predictive models' of the perceptually encountered world as minimizing complexity while maximizing accuracy. We note that other such integrations are possible; those who prefer their theoretical neuroscience in non-computational flavours may consult Allan Schore's (2012) synthesis. The value of Hopkins's chapter for this book, though, is to indicate another way in which psychoanalysis may seek to take its place amongst the sciences. As such it shows philosophy in its synoptic mode (see Gipps and Lacewing, 'Introduction: Know Thyself', this volume)—i.e. in its attempt to bring together disparate fields of knowledge with a philosophical eye to the maximal coherence of the proposed integration. We may note too the other scientific integrations now available and featuring in other sections of this

handbook—with anthropology (Narvaez, this volume), with sociology (Jay, this volume); or elsewhere—for example with cognitive and social psychology (Tesser 1991; Westen 2009).

Lacewing's chapter steps back from the scientific details to review the debate between those who would articulate psychoanalysis as science and those for whom it rather offers a hermeneutic. In fact that very distinction is one which—in the course of a careful unpicking of the assumptions at play in the debate of the last forty years between realist and constructionist strands in American psychoanalysis—is deconstructed in the course of his argument. That debate has, he contends, tended to confound its participants because it too often runs together matters which are more fruitfully kept separate. While it is possible to espouse all of realism (rather than constructionism), to respect (rather than deprecate) metapsychology, to believe in mental determinacy (rather than hold to the indeterminacy of the mental), and to portray psychoanalysis as aiming to provide scientific (as opposed to a different kind) of knowledge, these need not stand or fall together and should not in any case be assimilated. Although he is here not primarily making a case for a particular combination of these themes, one possibility that becomes particularly clear over the course of Lacewing's chapter is that of a realist hermeneutic science which respects mental indeterminacy while deprecating metapsychological speculation.

Matters take a somewhat different tack in the chapters by Fuchs and Gipps. If we understand what it is to be scientific in terms of *testability*, these chapters can be understood as highlighting the importance of primarily *non*-scientific aspects of psychoanalytic knowledge. That is, they are interested instead in our most basic *disclosure or characterization* of unconscious life—psychoanalysis as phenomenology. We do not test the aptness of a fundamental conceptual metaphor, picture, or vocabulary of unconscious life against its object, since that procedure presupposes that we have what here we do not have, namely some other conceptual scheme disclosing that object within which the comparison could be framed. Instead we consider its fruitfulness both in stimulating the development of further illuminations, and in warding off irremediably speculative, esoteric, or otherwise fruitless conceptions of our unconscious life.

Gipps aims to draw out how psychoanalysis provides a phenomenology in its articulations of unconscious life. The first half of his chapter makes a series of distinctions between our acknowledgement of and accounting for phenomena, grammar versus fact, revelation versus representation, and poiesis versus posits, and describes the central aspects of psychoanalytic theory in terms of the former element within each pair. The second half considers how a phenomenological understanding of psychoanalytic theory sheds light on the divergence between the (typical) psychoanalyst's and the (typical) psychologist's understanding of psychoanalytic theory and therapy.

Fuchs canvases several alternative phenomenological schemes for framing the dynamic unconscious, and to replace or at least complement the 'depth' metaphors of traditional psychoanalysis, he develops in detail a conception which puts to use Merleau-Ponty's notion of the lived body and Lewin's notion of a life space. For Fuchs and Merleau-Ponty,

the unconscious is sedimented in such bodily habits as we are blind to precisely because they shape our very experience itself.

Between them, these five chapters offer the basis for future work on the philosophical understanding of psychoanalysis. They analyse and deconstruct both psychoanalytic 'theory' and practice, and provide a range of approaches from which the question 'what is psychoanalysis?' may be answered. They seek to identify and preserve what is most valuable about psychoanalytic thought even while urging its development and integration with other philosophical and empirical approaches to understanding human beings.

References

Cioffi, F. (1999). *Freud and the Question of Pseudo-Science*. Chicago, IL: Open Court.

Freud, S. (1933). 'The question of a Weltanschauung'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 22 (pp. 158–182). London: Hogarth Press.

Habermas, J. (1987). *Knowledge and Human Interests*. Boston, MA: Polity Press.

Schore, A. (2012). *The Science of the Art of Psychotherapy*. New York: W. W. Norton.

Tesser, A. (1991). 'Social vs clinical approaches to self psychology: The self evaluation maintenance model and Kohutian object relations theory'. In R. Curtis (ed.), *The Relational Self: Theoretical Convergences in Psychoanalysis and Social Psychology* (pp. 257–281). New York: Guilford.

Westen, D. (2009). 'The cognitive self and the psychoanalytic self: can we put our selves together?', *Psychological Inquiry 3*: 1–13.

Richard G.T. Gipps

Dr Richard Gipps is a clinical psychologist in private psychotherapy practice in Oxford, UK, and an associate of the Philosophy Faculty at the University of Oxford. He convenes the Philosophy Special Interest Group of the Institute of Psychoanalysis, the Oxford Interdisciplinary Seminars in Psychoanalysis, and the Making the Unconscious Conscious seminar series. His research interests lie in psychoanalysis, psychosis, existential phenomenology, and Wittgenstein.

Michael Lacewing

Dr Michael Lacewing is a former Vice-Principal Academic and Reader in Philosophy at Heythrop College, London, an Honorary Reader in Research Department of Clinical, Educational and Health Psychology at University College, London, and a teacher of philosophy and theology at Christ's Hospital School, Sussex. He has published widely in philosophy of psychoanalysis, metaethics and moral psychology,

alongside writing textbooks for A level philosophy and training in Philosophy for Children (P4C).

Oxford Handbooks Online

Complexities in the Evaluation of the Scientific Status of Psychoanalysis 🔓

Morris N. Eagle

The Oxford Handbook of Philosophy and Psychoanalysis
Edited by Richard G.T. Gipps and Michael Lacewing

Abstract and Keywords

This chapter is primarily concerned with the scientific status of psychoanalysis as a theory of psychological functioning rather than as a form of treatment. It argues that certain characteristics of psychoanalysis, including theoretical formulations that are often obscure and untethered to observation, the proliferation of psychoanalytic 'schools', and habits of mind inculcated in psychoanalytic training and education count against according scientific status to psychoanalysis. The chapter also notes that psychoanalytic formulations have generated a good deal of research on psychological functioning and have yielded important insights into the nature of mind. It suggests that psychoanalytic education and training needs to change in ways that succeed in attenuating loyalty to particular psychoanalytic 'schools' and that encourage openness to empirical evidence and research and the operation of other self-corrective processes.

Keywords: scientific status, psychoanalytic training, clinical data, empirical evidence, theory of mind.

Morris N. Eagle

Introduction

IN addressing the question of the scientific status of psychoanalysis one needs to keep in mind: one, that psychoanalysis is a theory of psychological functioning and development as well as a form of treatment, each aspect generating different issues and questions; and two, that there is not a single psychoanalytic theory, but rather a collection of different psychoanalytic theories and 'schools'. The scientific status of Freudian theory has been debated at great length over a period of many years (e.g. Grunbaum 1984; Hopkins 1988; Lacewing 2018; Rosenblatt 1989; Sachs 1989). Far less discussed are the questions of: one, the scientific status of post-Freudian psychoanalytic theories (however, see Eagle

1986), formulations that have varying degrees of logical relationship to each other; and two, the degree to which self-corrective processes relating evidence to theory are embodied in the attitudes and practices of psychoanalytic practitioners (Cioffi 1970; Eagle 2014).

Although I comment on psychoanalytic psychotherapy process and outcome research, my main focus in this chapter will be on psychoanalysis as a theory of mind and psychological functioning rather than as a form of treatment. I think Freud was right when he stated that the claim of psychoanalysis on posterity will lie in its theory of mind rather than in its form of therapeutic practice.

(p. 354) Psychotherapy and Outcome Research

Like any therapeutic approach, the practice of psychoanalysis and psychoanalytic psychotherapy may have little to do with science, but may be largely a matter of craft and of individual talents and characteristics of the therapist, a conclusion supported by the evidence that individual differences among therapists account for far more of the variance in therapeutic outcome than specific techniques (Lambert and Barley 2001). Scientific-empirical questions enter the picture when one directs systematic inquiry into therapeutic outcome and the processes through which the outcomes occur.

There is a cadre of researchers (generally associated with university departments) who have carried out process and outcome research on psychoanalytic and psychodynamic treatment and have identified some of the factors that mediate therapeutic outcome. This focus on accountability is a great advance over reporting outcomes in the form of clinical anecdotes, self-selected clinical vignettes, testimonials, and 'clinical experience'.[1, 2] Thus, if this chapter dealt with the question of whether psychoanalytically oriented researchers are carrying out outcome and process studies on psychodynamic and psychoanalytic psychotherapy, the answer would be a resoundingly positive one. That is not where the problems related to the scientific status of psychoanalysis lie. The problems lie in the facts that one, to the extent that there is any interest on the part of a large part of the psychoanalytic community in such research, it is mainly limited to the demonstration that psychoanalytic psychotherapy 'works'; two, for the most part, there is little or no feedback from the research findings to psychoanalytic training, education, and clinical practice; and three, there is widespread hostility towards empirical research in the psychoanalytic community. Indeed, the very idea of carrying out research on psychoanalytic treatment has often met with indifference or hostility in the psychoanalytic community.[3] There are good reasons for scepticism regarding the direct value of psychotherapy research for clinical work with individual patients. Insofar as research studies often report group findings of a probabilistic nature, it would be mistaken and foolish to think that one can directly and simply apply such findings to the individual case. (p. 355) However, being aware of research findings can inform the background of one's clinical work and can lead one to question one's cherished and unquestioned assumptions.

Complexities in the Evaluation of the Scientific Status of Psychoanalysis

Let me provide an example. There is a significant body of psychoanalytic thought that views transference interpretations as virtually the sole useful interventions in psychoanalytic treatment (e.g. Gill 1982, 1994). However, there is evidence that transference interpretations can have negative effects on therapeutic outcome (e.g. Ryum et al. 2010; Schut et al. 2005) and that transference interpretations are relatively ineffective with patients who are functioning at a relatively high level of interpersonal relations and are effective mainly with patients who show relatively poor interpersonal relations (Hoglend et al. 2006, 2011; Johansson et al. 2010). The latter finding, in particular, contradicts the received wisdom that transference interpretations are likely to be more effective with higher functioning patients. Surely, any responsible clinician would want to be aware of such findings. Further, such findings are more likely to be useful to therapists than general data on the effect of number of transference interpretations on therapeutic outcome. Moreover, the further investigation of the factors involved in accounting for why patients with relatively poor object relations find transference interpretations more helpful, or which subset of patients within the poor object-relations category find transference interpretations especially helpful, would be even more likely to be useful in clinical practice. The more one includes individual difference variables in one's research, the less sharp the division between nomothetic and ideographic approaches to knowledge, and the greater the direct usefulness of research findings in clinical practice. The general point here is that rather than concluding that research findings are of no relevance to the individual clinician, findings generated by relatively sophisticated research designs can indeed be quite useful, even if their usefulness is limited to leading one to reflect on one's long-standing assumptions in clinical work.

The fact is, however, that most psychoanalytic institutes, particularly free-standing ones, neither require the carrying out of research nor, more noteworthily, prepare their candidates to become critical research consumers. Neither of these activities plays any significant role in psychoanalytic training and education in the majority of psychoanalytic institutes. Furthermore, although not true of all psychoanalysts, as noted, there is a pervasive atmosphere of hostility to research in most psychoanalytic institutes. As described by Kernberg (2012), one of the site visitors' criticisms of the Columbia University Psychoanalytic Institute—an institute more supportive of research—was that an emphasis on research would negatively affect a candidate's clinical training and work.

Kernberg's (2015) critique of the training and education practices of most psychoanalytic institutes echoes Glover's similar disaffection as far back as 1952. He writes:

> It is scarcely to be expected that a student who has spent some years under the artificial and sometimes hothouse conditions of a training analysis and whose professional career depends on overcoming 'resistance' to the satisfaction of his training analyst, can be in a favorable position to defend his scientific integrity (p. 356) against his analyst's theory and practice. And the longer he remains in training analysis, the less likely he is to do so. For according to his analyst the

candidate's objections to interpretations rate as 'resistances.' In short there is a tendency inherent in the training situation to perpetuate error.

(Glover 1952: 403)

Glover (1952) also describes the general state of affairs in the psychoanalytic community as follows:

An analyst, let us say, of established prestige and seniority, produces a paper advancing some new point of view or alleged discovery in the theoretical or clinical field. Given sufficient enthusiasm and persuasiveness, or even just plain dogmatism on the part of the author, the chances are that without any check, this view or alleged discovery will gain currency, will be quoted and re-quoted until it attains the status of an accepted conclusion. Some few observers who have been stimulated by the new idea may test it in their clinical practice. If they can corroborate it they will no doubt report the fact; but if they do not, or if they feel disposed to reject it, this scientific 'negative' is much less likely to be expressed, at any rate in public, and so, failing effective examination, the view is ultimately canonized with the sanctioning phrase 'so and so has shown'. In other words, an *ipse dixit* acquires the validity of an attested conclusion on hearsay evidence only.

(Glover 1952: 403)

Quite disturbingly, there appears to be little or no change in the practices and attitudes of the psychoanalytic community more than sixty years after Glover's comments were made.

The psychoanalytic community's interest in research appears to be directed mainly to findings that suggest the effectiveness of a psychodynamic approach. During the time that I was president of the Division of Psychoanalysis (Division 39) of the American Psychological Association in 1995–6 there was little interest in research among most of its members. A sudden interest in research arose in reaction to a spate of published criticisms launched against the lack of evidence for the effectiveness of psychoanalytic treatment. Some Division officers proposed the formation of a committee whose task it was to amass all the evidence available demonstrating the effectiveness of psychodynamic and psychoanalytic treatment. For the most part, the attitude towards research of individuals making this proposal generally ranged from indifference to hostility; they turned to it only for its demonstration and public relations value. They appeared to want to have it both ways, namely, take the position that psychotherapy research is irrelevant to psychoanalysis and, at the same time, turn to it for its public relations value. This attitude towards research on the part of Division 39 members continues to the present day: there is no longer a research section of the Division, partly due to lack of interest in research papers among its members.

To sum up, as far as a psychoanalytic ethos is concerned, the set of attitudes and habits of mind dominant among psychoanalysts and normative at psychoanalytic institutes is

resistant to systematic inquiry into therapeutic process and outcome. Such inquiry, when it does occur, is limited to a search for confirmation.

(p. 357) Psychoanalysis as Theory of Mind, Personality Development, and Psychopathology

Research on psychoanalytic theories and hypotheses has a long history going at least as far back as a study by Pötzl (1917) on the influence of 'indifferent impressions' on dream content and including, among other work, the Rosenzweig (1934) study on repression, the Sears (1943, 1944) studies on Freudian concepts at Stanford; the anthropological investigations of Malinowski (1927), and of Whiting (1960) and Whiting and Child (1953) at Yale; studies on defence and 'repressive style' (e.g. Weinberger 1990; Weinberger et al. 1979); research on ability to delay gratification (e.g. Mischel 1961); the early work on 'experimental neurosis' (e.g. Masserman 1944, 1950), reflecting an attempt to integrate Pavlovian and Freudian ideas (French 1933); research on Winnicott's (1953) concept of the transitional object (e.g. Cohen and Clark 1984); studies on the relationship between the Freudian notion of orality and dependency (e.g. Bornstein, Poynton, and Masling 1985); the Fisher and Greenberg (1996) volume of empirical test of psychoanalytic hypotheses; Masling and colleagues' volumes on empirical test of psychoanalytic theories (Bornstein and Masling 1998; Masling 1983, 1986, 1990; Masling and Bornstein 1993, 1994, 1996); and research on alexithymia (e.g. Taylor and Bagby 2013).

In short, psychoanalytic theory has been remarkably heuristic in generating a large body of empirical research—mainly by non-psychoanalysts outside the context of psychoanalytic training institutes. However: one, much of this research has had to do with Freudian theory, largely due to the susceptibility to empirical testing of at least a wide range of Freudian hypotheses; two, there appears to be a steady decrease outside psychotherapy research in the number of research papers on psychoanalytic concepts and formulations, largely due, I suggest later, to the unsusceptibility of many post-Freudian theoretical formulations to empirical research; and three, the large body of research findings have had little or no impact on psychoanalytic theorizing, perhaps because of the widespread insistence that only clinical data from the psychoanalytic treatment situation are relevant to psychoanalytic theorizing. I turn now to that assumption.

Clinical Data

A virtually canonical assumption in the psychoanalytic literature, cutting across different schools, is that evidence testing psychoanalytic formulations must be generated by the clinical situation. Although in certain contexts, data derived from the clinical situation can be useful in testing certain psychoanalytic hypotheses (see, for example, the research (p. 358) of Luborsky on 'momentary forgetting' (1967, 1973) and the research of Weiss and Sampson et al. (1986) on the relationship between therapist test-passing versus test-

failing and the emergence of warded off contents), as Grunbaum (1984) cogently argued, there are a number of problems with the uncritical use of such data. One problem is illustrated in a paper by Masling and Cohen (1987) in which they follow the course of a number of sessions of psychotherapy carried out by Carl Rogers, who defined his approach as non-directive. In the early sessions the patient's statements seem to be about equally divided between concern with issues of sex and self. Undoubtedly unwittingly, Rogers is silent when the patient talks about sex and tends to respond with a 'Mmm' or tends to look up when the patient talks about issues of self. Over time, the patient's statements about sex decrease and his statements about self increase. This vignette illustrates both the issue of suggestion in the clinical data as well as the problem of confirmation bias, the latter insofar as the patient's behaviour would tend to reinforce Rogers's belief that issues of self are more important than sexual issues for the patient.

Another problem with relying solely on clinical data to evaluate theoretical formulation is that patients' productions in the clinical situation are often ambiguous and subject to different interpretations. Given the analyst's theoretical commitment, he or she is likely to interpret the patient's free associations in accord with that commitment. For example, if one believes that infantile sexual and aggressive wishes culminating in the Oedipus complex are universal and at the core of psychopathology, one will likely find confirming 'evidence' for that belief in the patient's productions (see Peterfreund 1983). And similarly, if one believes, as Kohut (1984) does, that the need for empathic mirroring is universal and that a traumatic lack of empathic mirroring is at the core of psychopathology, one is likely to find confirming 'evidence' for that belief in the patient's productions. Confirmation bias is undoubtedly present in other disciplines. The important question, however, is the degree to which the discipline allows and encourages self-corrective practices (see Lacewing 2013).

An additional problem arises in the use of clinical data to defend etiological theories and hypotheses (e.g. early lack of empathic mirroring leads to lack of self-cohesiveness in adulthood). Such claims obviously require extra-clinical data. The shaky evidential base of psychoanalytic etiological formulations becomes apparent when one considers that virtually all such formulations are based on follow-back retrospective data rather than follow-up prospective data (see Kohlberg et al. 1972). Employing follow-back data, particularly follow-back clinical data, we develop our etiological theories of psychopathology based solely on patients seen in the clinical situation. We do not get data on individuals who may have been subject to the same vicissitudes as the patients we see, but who function quite adequately. This inevitably leads to an overestimation of the general pathogenic significance of the vicissitudes.

For example, in a follow-back study, Robins and McEvoy (1966) found that 59 per cent of adult drug users were truants as juveniles. However, employing follow-up or longitudinal data, they found that, for children who were truants before age fifteen, 26 per cent were drug users as adults. Hence, an etiological theory linking juvenile truancy to adult drug

use that is based solely on follow-back data would be seriously misleading regarding the strength of the association between the two factors.

(p. 359) Or consider this purely hypothetical example: We have data on 1,000 individuals subjected to sexual abuse as children. Ten per cent of the children sexually abused as children are diagnosed with borderline personality disorder (BPD) as young adults. The 100 young BPD adults will be far more likely to be seen clinically than the 900 others. Suppose further that all 100 of these young adults are seen at a particular clinic. The clinicians treating them there may then form an aetiological theory that significantly overestimates the impact of sexual abuse on the development of BPD. Researchers considering the longitudinal data on all 1,000 sexually abused individuals, however, will find that only 10 per cent of sexually abused children develop BPD, and so will conclude that although sexual abuse is, indeed, a risk factor, it must interact with other factors in determining the development of later BPD. Such research then encourages additional research to identify the factors that jointly precipitate BPD. In short, formulations based on clinical data on adult patients rely on follow-back or retrospective data rather than follow-up prospective data and therefore cannot generate adequate developmental theories. Recognizing this limitation and the importance of including methods and data from other disciplines is necessary if one is to develop an adequate psychoanalytic theory (or any other theory) of development, including the development of psychopathology.

Quite remarkably, for the most part, follow-back clinical data from adult patients constitute the primary basis for central psychoanalytic developmental theories, for example, Freud's (1917 [1916–17], 1925 [1924]) positing of a universal Oedipus complex that emerges at a particular age or, as another example, Kohut's (1984) assertion that lack of self-cohesiveness in adulthood is the consequence of early failures to have the infant's and child's needs for empathic mirroring met.

With regard to the former, even if unconscious incestuous wishes were found in all patients, this would hardly constitute adequate evidence for the positing of a universal Oedipus complex characterized by specific wishes and regularly emerging in all children at a particular age. Individuals seen in treatment, perhaps particularly in psychoanalytic treatment, may constitute a selective group different in important respects from other groups. Indeed, there is evidence that intergenerational conflicts (of which the Oedipus complex is an instance) are more likely to be seen in troubled families (Ascherman and Safier 1990; Erikson 1993; Flores, Mattos, and Salzano 1998). In short, an adequate evaluation of the Oedipus complex hypothesis cannot legitimately rest on clinical data, but rather requires longitudinal data with adequate sampling carried out outside the clinical context. Furthermore, the insistence that only clinical data are relevant to testing psychoanalytic hypotheses is but one instance of a stance that obstructs the possibility of refutation.

Not infrequently, one finds the claim in the psychoanalytic literature that clinical data from adult patients are not only an adequate, but the *only legitimate*, data for psychoanalytic developmental formulations. For example, Green (2000), an influential

spokesperson for this position, has distinguished between the 'real child' of empirical observation and the 'true child' of psychoanalysis and written that 'as far as psychoanalysis is concerned, the "real child" is not our concern' (2000: 61); only the 'true child' is relevant to psychoanalysis. What Green means by the 'true child' is 'the reconstructed child of psychoanalysis' (2000: 61). That is, Green seems to be saying that one can infer the (p. 360) course of development and mental states of the child on the basis of the productions of adult patients in psychoanalytic treatment.

If Green is saying that it is the patient's subjective sense of his or her childhood, that is, his or her psychic reality, rather than data from infant observation that are important in clinical work, his argument is well taken. This point, I suspect, would be apparent to anyone engaged in clinical work. Further, this claim does not require a sharp distinction between the 'true' child and the 'real' child. Nor, and this is a critical point, does it suggest that data from infant and child observation are irrelevant to psychoanalytic theories of the nature of psychological development. After all, psychoanalysis is a theory of mind as well as a clinical treatment.

Although Green's view can be interpreted as applying to the context of the clinical situation, the same cannot be said of Winnicott's (1960) statement that 'Indeed, it is not from direct observation of infants so much as it is from the study of the transference in the analytic setting that it is possible *to gain a clear view of what takes place in infancy itself*' (Winnicott 1960: 595, my emphasis). He goes on to say that 'Freud was able to discover infantile sexuality in a new way because he reconstructed it from his analytic work with his psycho-neurotic patients. In extending his work to cover the treatment of the borderline psychotic patient it is possible for us to reconstruct the dynamics of infancy and of infantile dependence, and of the maternal care that meets this dependence' (Winnicott 1960: 595).

While one can perhaps understand the usefulness of this approach in the context of treatment, it seems clear that one cannot legitimately base theories of infant and child development on data derived from the productions of adult patients in treatment—even if these data were not influenced by the therapist's theoretical commitment. If one wants to try to understand infant and child development, one observes infants and children in systematic ways and over long periods of time. It is remarkable that this is even a debated issue in the psychoanalytic community. That it is a debated issue, however, is important information in addressing the role of the habits of mind and attitudes of psychoanalytic theorists in the assessment of the scientific status of psychoanalysis. To sum up, the three problems with the uncritical use of clinical data for general theoretical formulations are: (1) suggestion; (2) confirmation bias; and (3) the inadequacy of follow-back data for etiological formulations.

It is important to note that clinical data can be very useful in formulating and testing a variety of clinical hypotheses and formulations. As noted earlier, this is illustrated in the work of Luborsky (1967, 1973) on 'momentary forgetting' and the research of Weiss and Sampson et al. (1986) on the relationship between therapist 'test-passing' and 'test-

failing' and the emergence of warded off contents (see also Silberschatz 2005). In these studies, the authors rely on a quasi-experimental design that is appropriate to the phenomena being investigated. For example, Luborsky (1967, 1973) found that compared to 'control' segments of clinical sessions in which momentary forgetting was not present, those segments of clinical sessions that preceded momentary forgetting were independently judged to be characterized by conflict and anxiety, suggesting that what one might (p. 361) refer to as 'mini-repressions' occur when conflictual and anxiety-laden material is being dealt with. With regard to the Weiss and Sampson researchers, based on the clinical data, compared to independent judgements of therapist test-failing (i.e. the independent judgement that therapist interventions tended to confirm the patient's unconscious pathogenic beliefs), therapist test-passing (i.e. the judgement that pathogenic beliefs are disconfirmed by the interventions) was reliably associated with the emergence in the clinical sessions of warded-off contents, broadened interests, and reduced anxiety.

Case Studies

Another means by which clinical data can constitute a legitimate means of gaining knowledge and a legitimate method of empirical research is the use of disciplined case studies (Kazdin 1981, 2008). However, as noted by Portuges (1994),[4] the American Psychoanalytic Association Committee headed by Klumpner and Frank (1991) reviewed the sixty most frequently referenced 'psychoanalytic articles' that had been published in three prominent journals between 1869 and 1982 and scrutinized the top fifteen. They reported:

> [n]ot a single one of these 15 papers included any significant amount of ordinary clinical data. Our sample contained no case studies. What evidence was given in support of the conclusions could only be described as sketchy. We found no verbatim examples of and only one dream fragment … Our most frequent descriptions of the 15 papers were, 'No clinical data', 'No evidence', 'Overgeneralized' or 'Assumptions not testable'. In short, it was as if psychoanalytic data were considered so well documented that the reader would be interested only in questions of interpretation.(as cited in Portuges 1994: 12)

There have been attempts to identify and recommend the standards for the case study method that would enhance its reliability and validity as data (e.g. Eagle and Wolitzky 2011; Edelson 1984). However, as is the case with other empirical data, these attempts have had little impact on papers published in psychoanalytic journals. Little has changed since the Klumpner and Frank (1991) report. The clinical data reported in most psychoanalytic papers tend to be limited to self-selected and sketchy clinical vignettes.

(p. 362) Untestability of Formulations

Whatever the limitations of the above kind of theorizing, for the most part the reader can generally understand what formulations are being proposed and perhaps imagine the kind of evidence that might be relevant to assessing these formulations. Unfortunately the psychoanalytic literature is replete with formulations that are so vague and/or so obscure that it is difficult to conceive of any evidence that could play a role in their evaluation. Especially unfortunate is the fact that the authors of these formulations are not obscure figures, but include the most prominent and influential psychoanalytic theorists in the world today.

Consider, for example, the following passage which is representative of Klein's formulations of infant fantasies. Klein (1930: 24–5) writes:

> In my experience sadism reaches its zenith in this [early] phase, which is ushered in by the oral-sadistic desire to devour the mother's breast (or the mother herself) … It is my experience that in the phantasied attack on the mother's body a considerable part is played by the urethral and anal sadism which is very soon added to the oral and muscular sadism. In phantasy the excreta are transformed into dangerous weapons: wetting is regarded as cutting, stabbing, burning, drowning, while the faecal mass is equated with weapons and missiles.

> (Klein 1930: 24–5)

Other than the comment 'It is my experience', Klein offers no evidential basis for the attribution of such florid fantasies to the infant. Indeed, as I have noted, it is difficult to even imagine what kind of clinical evidence could serve as a basis for such attributions. Whereas one can think of evidence that could support Freud's positing of a universal Oedipal stage or Kohut's claim regarding the need for empathic mirroring, it is very difficult to even imagine the kind of evidence that would support Klein's description of the infant's experience. This leaves one with the question of how Klein knows about the experiences of the young infant. It also raises the question of why such theory should be taken seriously, let alone taught at psychoanalytic institutes.

The writings, particularly the late work, of Bion, another highly influential psychoanalytic theorist, fare no better with regard to the issue of testability. In his later writings, Bion became increasingly mystical and preoccupied with the concept of 'O', which he variously defined as the 'Absolute Truth', 'Ultimate Reality', 'infinity', 'noumena or things in themselves', and 'godhead' (Bion 1965: 147–9; 1970: 52). Even those sympathetic to Bion's thought have commented on the problems with his later work. For example, O'Shaughnessy (2005) writes in regard to this work:

> Contradictions have their appeal; breaking the laws of thought and reason bring a quantum of verbal fun. Yet, in scientific writings, such transgressions lead us to anything and everything we fancy—because, as is readily logically demonstrable,

from a contradiction any proposition follows. Texts with contradictions risk an unending proliferation of meanings.

(O'Shaughnessy 2005: 1524–5)

(p. 363) The exegesis of Bion's writings does not help much. For example, in a recent paper sympathetic to Bion's ideas, Rosegrant (2012) refers to the notorious difficulty of Bion's writing style and comments that 'a style that so dominates the message is better taken seriously in itself, not as incidental, but as an integral part of Bion's message' (Rosegrant 2012: 724). He then goes on to describe 'the psychotic mechanisms in Bion's writing' (724). These include 'attacks on linking' and 'denudation of meaning' (725), a major 'massive' example of which is Bion's notorious grid. Rosegrant writes with regard to the grid: 'he may or may not be describing psychologically and clinically useful ideas, but displaying these ideas in a grid is an attack on the usually linked ideas that a grid will simplify and help with understanding' (727).

Rosengrant writes that 'Bion's writing, which on first glance appears to be about psychosis, is on a more fundamental level an inducement to psychosis. Without providing the grounding that would come from alerting the reader that this is about to happen, Bion immerses the reader in the experience he is ostensibly describing: to read Bion is to be psychotic' (Rosegrant 2012: 727). He then goes on to state: 'to recognize that Bion's writing is mad does not mean that it is without value or meaning, but clarifies that its value and meaning are romantic' (727–8). Rosegrant makes clear that what he means by romantic in the present context is essentially a search for the 'mystical one, god, the god head' (729) that Bion refers to as O. Further according to Rosegrant, for Bion, 'the experience of O is the proper goal of analysis' (730).

As other examples of exegeses of Bion's work, consider the following passages from a prominent Kleinian and Bionian psychoanalyst's account of Bion's ideas:

1. 'Bion took the final step in bringing psychoanalysis into alignment with virtually all the fields of knowledge, including mathematics, physics, aesthetics, literature, mythology, philosophy, poetry, religion, archeology, etc.—all for the purpose of putting it on a veritable Internet of a universal epistemology' (Grotstein 1997: 78).
2. 'The major themes that express this profound paradigm shift include the concept and the penumbra of ideas that cluster around "O", the concept of the mystic, the genius, and the "messiah", the invocation of the "religious instinct", the very idea of God and the Godhead, and his explorations of fetal mental life' (Grotstein 1997: 78).
3. 'It is my impression upon reading Bion that this epistemological-ontological-phenomenological cycle takes place as follows: the infantile portion of the personality experiences O, Absolute Truth, as beta-elements, projects them into the analyst who, in containing them, transforms them from O to K (from Truth to Knowledge) and shares this knowledge with the analysand, who, upon this acceptance of the *knowable truth* allows it to become transformed into wisdom, personal O. I term this last act the achievement of the transcendent position' (Grotstein 2004: 1086).

Klein's and Bion's concepts and formulations may be meaningful, evocative, and helpful to many psychoanalytic clinicians. However, that is not the point here. Insofar as this (p. 364) chapter is concerned with the scientific status of psychoanalytic theories, the issues of evidence and empirical testability are, perforce, central. Further, insofar as Klein and Bion are not marginal figures, but rather two of the most prominent and influential theorists in the psychoanalytic world today, with a large group of followers, the fact that many of their formulations are not susceptible to empirical test is a significant factor in evaluating the status of psychoanalytic theory. This judgement, however, does not preclude the value of their work in other contexts.

Progress and Theory Change

From the earliest period of its history to the present day, the emergence of 'isms' and 'schools', including Jungian, Adlerian, Horneyian, Sullivanian, Kleinian, Bionian, Lacanian, ego psychology, modern Freudian, Fairbairnian, and relational psychoanalytic theories, has characterized theoretical developments in psychoanalysis. Furthermore, the emergence of each of these schools is not based on the presentation of evidence that refutes earlier theories or compelling new evidence that supports the new theory. Rather, theory change in psychoanalysis is characterized by the emergence of a charismatic figure who attracts loyal adherents, often culminating in the establishment of psychoanalytic institutes that provide training from a particular theoretical perspective (Eagle 1987, 1995; Eagle and Wolitzky 1989).

Gedo (1984) has noted, accurately, I believe, that 'experienced clinicians will refuse to alter their convictions on the bases of [the researchers'] results. Instead, they would continue to form their psychoanalytic views in direct response to their personal experiences … the belief that psychoanalysis will make progress by validating the best hypotheses through refined scientific methods is implausible. What we need are innovative ideas powerful enough to compel acceptance by significant portions of the analytic community' (Gedo 1984: 514). However liberal one's criteria for scientificity, it is difficult to think of a discipline as scientific where theory changes take place as a function not of systematic evidence but rather of the evocative power and appeal of its formulations and the charisma of the originators of these formulations.

The historical trajectory of psychoanalytic theorizing is not encouraging to anyone who looks for cumulative progress of a scientific sort. Indeed, although in certain areas one may discern progress, many theoretical developments can be characterized as retrogressive in one important sense. That is, whatever the problems and difficulties of Freudian theory, many of its propositions are formulated with sufficient clarity that they are susceptible to empirical test, as evidenced by the fact that much of the empirical research on psychoanalytic propositions is linked to Freudian theory. Contrastingly, as we have seen, Kleinian and neo-Kleinian theories, the most influential current psychoanalytic theories, are formulated in such a way that the possibility of scientific testing does not

arise. In so far as we are concerned with the development of a science of psychoanalysis, this is not an encouraging trajectory.

(p. 365) Falsifiability of Psychoanalytic Propositions

It has been argued that the possibility of falsifiability speaks to the scientificity of a theory, including psychoanalytic theory (Popper 2003/1965; See also Lakatos 1970). For example, arguing against Cioffi's (1970) position that the attitudes and practices of the psychoanalytic community that protect these propositions from falsifiability compromise its scientific status, Farrell (1961, 1984) has proposed that because certain specific psychoanalytic generalizations can be refuted in principle, psychoanalytic theory should be accorded scientific status. There are a number of problems with Farrell's proposal. One problem, already discussed, is that certain core psychoanalytic propositions do not appear to be susceptible to empirical refutation. For example, how would one go about devising procedures that would allow the possibility of refuting Freud's and Klein's claims regarding an inborn death instinct, let alone its influence on psychological development and functioning? The same question can be addressed to the florid fantasies attributed to young infants by Klein or to Bion's claims regarding beta and alpha elements and alpha functioning.

A second problem with Farrell's proposal is that, as Cioffi (1970) has pointed out, 'refutability in principle is not an adequate criterion of the genuinely empirical character of an enterprise' (Cioffi 1970: 472). Were it an adequate criterion for scientific status, astrology, for example, would have to be seen as scientific. For surely, astrological claims are not only refutable, but, indeed, have been refuted. The issue is, as Cioffi (1970) has noted, that astrological 'theorists' not only do not employ procedures that are able to test their claims but, indeed, engage in practices that obstruct refutation. Thus, if being refutable in principle is not accompanied by being refutable in practice, due to a parade of obstructional procedures of the defenders of the claim, refutability in principle becomes an empty criterion. The situation is somewhat analogous to the relationship between the law and the carrying out of the law. Citing Orwell, Remnick (2016) observes that the law itself affords no protection insofar as how laws are implemented depends on the individuals carrying them out and on 'the general temper of the country' (Remnick 2016: 2). We all know of nations which have freedom-affirming constitutions, but dictatorial practices.

Farrell's (1961) specific example of a psychoanalytic proposition that is refutable in principle is the theory of the Oedipus complex. Indeed, not only is there no convincing evidence in support of a universal Oedipus complex, but a good deal of evidence against that claim (see Eagle 2018). However, the refutation of the theory of the Oedipus complex has not resulted in its relinquishment—just as refutations of astrological claims have not resulted in their abandonment.

Complexities in the Evaluation of the Scientific Status of Psychoanalysis

Like Cioffi, Meehl (1993) also cites the practices and habits of mind of members of the psychoanalytic community when he writes: 'One need not conduct a literature search to realize that the divergences in theory and technique among therapists in a broadly (p. 366) "psychoanalytic" tradition are vast, increasing, and show little or no signs of the sort of cumulative self-corrective development characteristic of post-Galilean science' (Meehl 1993: 324). That is, there is no 'self-corrective' feedback from the evidence to the theory. The evidence contradicting the claim of a universal Oedipal stage of development is largely ignored by many Freudian theorists who continue to refer to the Oedipal complex both in their theoretical and clinical papers.

Another reaction to cogent critiques of the Oedipus complex hypothesis has been to retain the concept, but radically alter its meaning so that it no longer bears much resemblance to the original formulation and theoretical claims. Loewald (1979) writes: 'the oedipal attachments, struggles, and conflicts must also be understood as new versions of the basic union-individuation dilemma' (775). In describing 'oedipal attachments' this way, Loewald (1979) essentially redefines the Oedipus complex so that it is understood as entailing conflict between remaining tied to early parental figures versus separation and autonomy (in Mahler's [1968] terms, between symbiosis versus separation-individuation). The existence of unconscious incestuous and death wishes, the defining features of the Oedipus complex, is no longer posited. Rather, reference to such wishes is now understood not as pointing to actual incestuous and death wishes, but rather as *symbolic* expressions of union and separation-autonomy. However, despite gutting its essential features, which include the universality of unconscious incestuous and death wishes, Loewald continues to refer to the Oedipus complex.

As another example of redefinitions of the Oedipus complex, Morehead (1999) writes that although 'human and animal evidence supports the Westermarck hypothesis' (the hypothesis that prolonged propinquity dampens sexual interest), such evidence does not contradict 'recent analytic theory on the Oedipus Complex [which] does not require the existence of a central, powerful, incestuous sexual drive' (347). Like Loewald, Morehead removes a core aspect of the Oedipus complex—the existence of a central powerful sexual drive—but nevertheless insists that it remains a viable theory. In short, both Loewald and Morehead illustrate obstructional attempts to preserve a central psychoanalytic hypothesis through nominal means despite the dearth of confirming evidence and the accumulation of negative evidence.[5]

One could perhaps argue that the issue is largely a verbal one, i.e. that Loewald and Morehead have merely retained the term 'Oedipus complex' to refer to certain dynamics. However, this argument is not convincing. For one thing, Morehead makes clear that his intention is to *preserve* the theory of the Oedipus complex. Perhaps most important, Loewald's and Morehead's manoeuvres obscure the relationship between evidence and theory through nominal means. That is, they obscure the conclusion that there is little or no evidence in support of Freud's positing of universal incestuous and death wishes. Nor is there evidence that how one resolves Oedipal wishes has a decisive influence on psychological development, including development of gender identity and the structure of

the superego. Given this state of affairs, the warranted conclusion would point to the relinquishment of the Freudian theory of the Oedipus complex. Further, (p. 367) there *is* much evidence that issues of union versus separation-individuation and how they are dealt with *do* exert a major influence on psychological development. However, union versus separation-individuation is not a restatement or reconceptualization of the Oedipus complex but, in an important sense, is a *replacement* of Oedipal theory.

Truth, Science, and a Heuristic Discipline

Despite the difficulties identified earlier, it is clear that psychoanalytic theory has had much to offer to those interested in how the mind works. It has attracted the interest of Nobel laureates such as Adrian (1946) and Kandel and gifted scientists such as Gallese, the co-discoverer of mirror neurons. Indeed, both Kandel (2007) and Gallese (personal communication) have reported that at one time in their life, they seriously thought of becoming a psychoanalyst. It is important to add, however, that the inspiration for these scientists was the corpus of Freud's work on the nature of the mind rather than the post-Freudian theorizing I have described.

Despite the problems I have discussed, there have been post-Freudian theoretical developments that have contributed to an understanding of mental functioning. In my view, these contributions include certain aspects of ego psychology (e.g. the importance of certain ego functions such as reflective capacity, ability to delay gratification, and capacity for affect regulation); the importance of what Mahler (1968) refers to as separation-individuation in development; and Bowlby's (1969, 1973, 1980) attachment theory, particularly its account of the basis for the infant–mother bond and the developmental consequences of the particular nature of that bond.

Although much of this work may have been stimulated by psychoanalytic ideas, for the most part research in these areas takes place outside the psychoanalytic context (another reason that psychoanalysts should familiarize themselves with research relevant to psychoanalysis). For example, research and theory on ego functions (although it is not referred to that way in the non-psychoanalytic literature) is primarily cognitive psychology. Similarly, although Bowlby was trained as a psychoanalyst and viewed himself as one throughout his life, virtually all the burgeoning research and theoretical developments in attachment theory have taken place outside psychoanalysis, in domains such as developmental psychology. These examples suggest that many psychoanalytic ideas, including critical reactions to them, have been richly heuristic in stimulating research. That is, to borrow Reichenbach's (1938/1951) admittedly problematic distinction, one can say that many psychoanalytic ideas have been useful in the context of discovery; the problems lie in the context of testing them.

The value of a body of thought is not defined by its scientific status, nor is it the case that only scientific method can generate important insights and truths. I believe that, at its best, a psychoanalytic mode of thought is of great value in a number of ways, and in particular, it may provide important insights into experience and behaviour that are not

(p. 368) readily available through other means. As an example of this phenomenon, I am reminded of a description of an experience provided by Meehl (1984). He describes observing a man and woman leaving a hospital, the woman carrying a Raggedy Ann doll. Meehl becomes unbearably sad and begins sobbing, imagining that the man and woman have just left their very ill child. Upon arriving at his analytic session, still sobbing, his analyst asks Meehl whether he was mean to his daughter before coming to the session, upon which Meehl becomes immediately aware of the reasons for his sadness and sobbing, which then completely cease.

Let me provide another example from my own clinical experience. Many years ago, I was working with a patient, A. G., whose presenting symptom was his obsessive fear that he might be homosexual, which, if true, would represent for him an abominable sin. He was tormented by these thoughts. The symptom appeared after his girlfriend pressed for them to become engaged, with an eye towards marriage. During one session, he reported that after dinner with his mother and father (at the time, he was still living at home), when his mother asked him to put a baking dish back on a high shelf (he was quite tall), his obsessive thoughts about whether he was homosexual got worse. When I asked him whether he had any further thoughts about this incident, he said that putting the baking dish away was something his father used to do. He then remembered that when his mother had recently asked him to mow the lawn—remarking again that this was something his father used to do—the same thing occurred. He then volunteered that after he received a job promotion, his obsessive thoughts reached their peak in frequency and intensity.

Although other interpretations are certainly possible, given other information about A. G.'s experiences and difficulties, a plausible hypothesis close to the clinical evidence is that a central symbolic meaning of all three events—putting the baking dish away, mowing the lawn, and receiving a job promotion—that were associated with an exacerbation of A. G.'s symptoms is that they all involved being placed in an adult role and confronting the challenges and responsibilities that that entails—a role in which he does things his father used to do.

And, indeed, A. G.'s severe anxiety about leaving home, both figuratively and literally, and feeling panicky and engulfed at the thought of marriage, emerged as the central issues in treatment. The point I want to make with this clinical anecdote, similar to Meehl's point about his own experience, is the following one: I know of no theory other than psychoanalytic theory, particularly its formulation of primary process thinking that (in the case of A. G.) enables one to identify the common symbolic meaning among activities as disparate as placing a baking dish on a high shelf, mowing the lawn, and receiving a job promotion. I also do not know of any other theory that provides as clear an insight into the defensive function of a troubling symptom.[6]

(p. 369) Ironically, this very strength of psychoanalytic theory in finding hidden common motives and meanings in disparate activities is also a potential problem. This is so because the meanings identified in the patient's material are often arbitrary, far removed

from the patient's experience, and vary with the analyst's theoretical orientation (see Peterfreund 1983). Therefore, it becomes very difficult to distinguish between arbitrary theoretical impositions on to the clinical material and interpretations warranted by the clinical evidence. As Meehl (1995) puts it, 'the clear message of history is that anecdotal method delivers both wheat and chaff, but it does not enable us to tell which is which' (1019). Being better able to make this distinction would represent a significant contribution. I think the earlier noted work of psychodynamically oriented researchers such as Luborsky and Weiss and Sampson gives us some idea of how such distinctions can be made. However, large sections of the psychoanalytic community need to overcome their indifference or hostility towards these sorts of studies.

Conclusions

1) Although, as Freud and others believed, the development of psychoanalysis as a separate discipline with its own free-standing training institutes and its relative isolation from other disciplines may have made it possible for it to emerge as a coherent body of thought, its status as a separate discipline with its own education and training institutes has constituted a major barrier to theoretical (and perhaps clinical) progress. For the most part, psychoanalytic associations and training institutes have been more a force for the preservation of guild identity and orthodoxy (including pluralistic orthodoxies) than a means for theoretical and clinical progress (Kernberg 2012, 2015).

2) There appears to be a more recent acceptance of empirical research and greater valuing of the importance of establishing links between psychoanalysis and other disciplines in some quarters of the psychoanalytic community. However, this must be broadened beyond a selective search for findings confirming psychoanalytic hypotheses to reflect an equal interest in findings that might *challenge* psychoanalytic formulations and serve as an impetus to theoretical change and progress.

3) In contrast to central aspects of Freudian theory which, at least in principle, lend themselves to empirical test, many post-Freudian theoretical developments, for example those associated with Klein and Bion, are often so arbitrary, obscure, and untethered to observation, that the question of testability becomes moot. Thus, the historical trajectory of psychoanalytic theorizing does not provide a basis for confidence in the prospect of steady and cumulative progress in psychoanalytic theorizing.

4) Theory change in psychoanalysis comes about not by virtue of new findings, but through the emergence of charismatic figures who attract loyal followers who establish psychoanalytic training institutes that promulgate the ideas of new 'schools'.

(p. 370) The general picture that emerges is that of two cultures, one, so to speak, within the psychoanalytic community of psychoanalytic institutes, and the other outside that community, most often associated with university departments. The latter group,

committed to a scientific ethos, carries out research on propositions directly generated or inspired by psychoanalytic theory (e.g. Blatt 2008 on self-definition and relatedness; Cramer 2008, 2012 on defence; Fonagy and Bateman 2006 on reflective capacity and mentalization). The former group, made up of the great majority of psychoanalysts, tends either to have no interest in or to be hostile to research on psychoanalytic formulations. Further, many of the formulations that are most influential in the psychoanalytic community are not susceptible to empirical test. And, perhaps most important, it is this community that shapes the education, training, and therefore, habits of mind of most psychoanalysts.

To sum up, it seems clear that, at its best, a psychoanalytic mode of thought can generate important insights regarding the nature of mind that may not be readily available through other means. However, at its worst, it has generated arbitrary and obfuscatory formulations far removed from observation and unsusceptible to empirical test by any means. Further, given the nature of psychoanalytic training and education, there appears to be no reliable self-correcting mechanism that enables psychoanalysts to distinguish between the plausible and testable, on the one hand, and the obscure and untestable, on the other. Indeed, if one looks at the historical trajectory of psychoanalytic theorizing from Freudian to dominant contemporary theories, it appears that at least as much influential psychoanalytic theorizing is increasingly characterized by obtuse formulation far removed from observation and evidence. Unhappily, this is a grim picture but, I think, an accurate one. I agree with Kernberg (2012, 2015) that necessary modifications in training and education may serve to brighten this picture. I believe that current practices and habits of mind regarding such matters as attitudes towards evidence and testability need to change. The challenge, in my view, is to be able to integrate the best of what a psychoanalytic mode of thought has to offer with the attenuation of loyalties to particular psychoanalytic 'schools' and an openness to empirical evidence and research in the service of self-corrective practices and procedures.

References

Adrian, E. D. (1946). 'The mental and physical origins of behavior'. *The International Journal of Psychoanalysis 27*: 1–6.

Ascherman, L. I. and Safier, E. J. (1990). 'Sibling incest: A consequence of individual and family dysfunction'. *Bulletin of the Menninger Clinic 54*(3): 311.

Berlin, H. A. (2011). 'The neural basis of the dynamic unconscious'. *Neuropsychoanalysis 13*(1): 5–31.

Bion, W. R. (1954). 'Notes on the theory of schizophrenia'. *The International Journal of Psychoanalysis 35*: 13–18.

Bion, W. R. (1965). *Transformations*. New York: Basic Books.

Bion, W. R. (1970). *Attention and Interpretation*. London: Tavistock.

Blatt, S. J. (2008). *Polarities of experience: Relatedness and self-definition in personality development, psychopathology, and the therapeutic Process*. Washington, DC: American Psychological Association Books.

(p. 371) Bornstein, R. F. and Masling, J. (eds) (1998). *Empirical perspectives on the psychoanalytic unconscious: Empirical studies of psychoanalytic theories*, Vol. 7. Washington, DC: American Psychological Association.

Bornstein, R. F., Poynton, F. G., and Masling, J. (1985). 'Orality and depression: An empirical study'. *Psychoanalytic Psychology 2*(3): 241–249.

Bowlby, J. (1969/1982). *Attachment and Loss*, Vol. 1, *Attachment*, 2nd edn. New York: Basic Books. (Original work published 1969).

Bowlby, J. (1973). *Attachment and Loss*, Vol. 2, *Separation*. New York: Basic Books.

Bowlby, J. (1980). *Attachment and Loss*, Vol. 3, *Loss: Sadness and Depression*. New York: Basic Books.

Cioffi, F. (1970). 'Freud and the idea of a pseudo-science'. In F. Cioffi and R. Borger (eds), *Explanation in the behavioral sciences* (pp. 471–499). Cambridge, MA: Cambridge University Press.

Cohen, K. N. and Clark, J. A. (1984). 'Transitional object attachments in early child and personality characteristics in later life'. *Journal of Personality and Social Psychology 46*(1): 106–111.

Cramer, P. (2008). 'Seven pillars of defense mechanisms theory'. *Social and Personality Psychology Compass 2*: 1963–1981.

Cramer, P. (2012). *The development of defense Mechanisms: Theory, Research, and Assessment*. New York: Springer.

Eagle, M. N. (1984). *Recent Developments in Psychoanalysis: A Critical Examination*. New York: McGraw-Hill.

Eagle, M. N. (1986). 'Critical notice: A. Grunbaum's "The Foundations of psychoanalysis" '. *Philosophy of Science 53*(1): 65–88.

Eagle, M. (1987). 'Theoretical and clinical shifts in psychoanalysis'. *American Journal of Orthopsychiatry* 57(2): 175–185.

Eagle, M. (1995). 'The dynamics of theory change in psychoanalysis'. In J. Earman, A. I. Janis, G. J. Massey, and N. Rescher (eds), *Philosophical Problems of the Internal and External World: Essays on the Philosophy of Adolf Grunbaum* (pp. 373–408). Pittsburg, PA: University of Pittsburg Press.

Eagle, M. N. (2011). 'Psychological processes and neural correlates'. *Neuropsychoanalysis 13*(1): 37–42.

Eagle, M. N. (2014). *The Scientific Status of Psychoanalysis*. Paper given at Conference on Psychoanalysis, Philosophy, Religion, and Art. London, October 2014.

Eagle, M. N. (2018a). *Core Concepts in Classical Psychoanalysis: Clinical, Research Evidence and Conceptual Critiques.* London and New York: Routledge.

Eagle, M. N. (2018b). *Core Concepts in Contemporary Psychoanalysis: Clinical, Research Evidence and Conceptual Critiques.* London and New York: Routledge.

Eagle, M. N. and Wolitzky, D. L. (1989). 'The idea of progress in psychoanalysis'. *Psychoanalysis and Contemporary Thought 12*(1): 27–72.

Eagle, M. N. and Wolitzky, D. L. (2011). 'Systematic empirical research versus clinical case studies: A valid antagonism? *Journal of the American Psychoanalytic Association 59*(4): 791–818.

Edelson, M. (1984). *Hypothesis and Evidence*. Chicago, IL: University of Chicago Press.

Ellenberger, H. (1970). *The Discovery of the Unconscious: The History and Evolution of Dynamic Psychiatry*. New York: Basic Books.

Erikson, M. (1993). 'Rethinking Oedipus: An evolutionary perspective of incest avoidance'. *American Journal of Psychiatry 150*: 411–416.

Farrell, B. A. (1961). 'Can psychoanalysis be refuted? *Inquiry 4*: 16–36.

Farrell, B. A. (1984). 'The standing of psychoanalysis'. *Nous 18*(3): 534–541.

(p. 372) Fenichel, O. (1945). *The Psychoanalytic Theory of Neurosis*. New York: W. W. Norton.

Firestein, S. (1978). *Termination in Psychoanalysis*. New York: International Universities Press.

Fisher, S. and Greenberg, R. (1996). *Freud Scientifically Reappraised: Testing the Theories and Therapy*. Oxford: John Wiley & Son.

Flores, R. Z., Mattos, L.F., and Salazano, F. M. (1998). 'Incest: Frequency, predisposing factor and effects in a Brazilian population'. *Current Anthropology 39*(4): 554–558.

Fonagy, P. and Bateman, A. (2006). 'Mechanisms of change in mentalization-based treatment of BPD'. *Journal of Clinical Psychology 62*(4): 411–430.

French, T. (1933). 'Interrelations between psychoanalysis and the experimental work of Pavlov'. *American Journal of Psychiatry 89*: 1165–1203.

Freud, S. (1917 [1916–17]). *Introductory Lectures in Psychoanalysis*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 16 (pp. 243–463). London: Hogarth Press.

Freud, S. (1925 [1924]). 'An autobiographical study'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 20 (pp. 1–74). London: Hogarth Press.

Gallese, V. (pers. comm.).

Gedo, J. (1984). 'Letter to the Editor'. *Review of Psychoanalytic Books 41*(4): 511–515.

Gill, M. M. (1982). *Analysis of Transference*, Vol. 1, *Theory and Technique*. New York: International Universities Press.

Gill, M. M. (1994). *Psychoanalysis in Transition*. Hillsdale, NJ: The Analytic Press.

Glover, E. (1952). 'Research methods in psychoanalysis'. *The International Journal of Psychoanalysis 33*: 403–409.

Goldberg, A. (1978). *The Psychology of the Self: A Casebook*. New York: International Universities Press.

Green, A. (2000). 'Science and science fiction in infant research'. In J. Sandler, A. M. Sandler, and R. Davies (eds), *Clinical and Observational Psychoanalytic Research* (pp. 41–73). London: Karnac Books.

Grotstein, J. S. (1997). 'Bion, the pariah of "O" '. *British Journal of Psychotherapy 14*(1): 77–90.

Grotstein, J. S. (2004). 'The seventh servant: The implication of a truth drive in Bion's theory of "O" '. *The International Journal of Psychoanalysis 85*: 1081–1101.

Grunbaum, A. (1984). *The Foundations of Psychoanalysis: A Philosophical Critique*. Berkeley/Los Angeles, CA: University of California Press.

Hoglend, P., Amlo, S., Marble, A., Bogwald, K. P., Sjaastad, M. C., and Heyerdahl, O. (2006). 'Analysis of the patient–therapist relationship in dynamic psychotherapy: An experimental study of transference interpretations'. *American Journal of Psychiatry 163*: 1739–1746.

Hoglend, P., Hersoug, A. G., Bogwald, K.-P., Amlo, S., Marble, A, Sorbye, O., and Crits-Christoph, P. (2011). 'Effects of transference work in the context of therapeutic alliance and quality of object relations'. *Journal of Consulting and Clinical Psychology 79*: 697–706.

Hopkins, J. (1988). 'Epistemology and depth psychology: Critical notes on the "Foundations of Psychoanalysis" '. In R. Clarke and C. Wright (eds), *Mind, Psychoanalysis, and Science* (pp. 33–60). Oxford: Blackwell.

Johansson, P., Høglend, P., Ulberg, R., Amlo, S., Marble, A., Bogwald, K., and Heyerdahl, O. (2010). 'The mediating role of insight for long-term improvements in psychodynamic therapy'. *Journal of Consulting and Clinical Psychology 78*: 438–448.

(p. 373) Kandel, E. (2007). *In Search of Memory: The Emergence of a New Science of Mind*. New York: W. W. Norton.

Kazdin, A. E. (1981). 'Drawing valid inferences from case studies'. *Journal of Consulting and Clinical Psychology 49*(2): 183–192.

Kazdin, A. E. (2008). 'Evidence-based treatment and practice: New opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care'. *American Psychologist 63*(3): 146–159.

Kernberg, O. F. (2012). 'Suicide prevention for psychoanalytic institutes and societies'. *Journal of the American Psychoanalytic Association 60*: 707–719.

Kernberg, O. F. (2015). 'Resistance and progress in developing research framework in psychoanalytic institutes'. *Psychoanalytic Inquiry 358*: 98–114.

Klein, M. (1930). 'The importance of symbol formation in the development of the ego'. *The International Journal of Psychoanalysis 11*: 24–39.

Klumpner, G. H. and Frank, A. (1991). 'On methods of reporting clinical material'. *Journal of the American Psychoanalytic Association 39*: 459–477.

Kohlberg, L., LaCrosse, G., and Ricks, D. (1972). 'The predictability of adult mental health from childhood behavior'. In B. B. Wolman (ed.), *Manual of Child Psychopathology*. New York: McGraw-Hill.

Kohut, H. (1984). *How Does Analysis Cure?* Chicago, IL: University of Chicago Press.

Lacewing, M. (2013). 'The problem of suggestion: an analysis and solution'. *Philosophical Psychology 26*: 718–743.

Lacewing, M. (2018). 'The science of psychoanalysis'. *Philosophy, Psychiatry, and Psychology 25*(2): 95–111.

Lakatos, I. (1970). 'Falsification and the methodology of scientific research programs'. In I. Lakatos and A. Musgrave (eds), *Criticism and the Growth of Knowledge* (pp. 91–95). New York: Cambridge University Press.

Lambert, M. J. and Barley, D. E. (2001). 'Research summary on the therapeutic relationship and psychotherapy outcome'. *Psychotherapy: Theory, Research, Practice, Training 38* (4): 357–361.

Loewald, H. (1979). 'The waning of the Oedipal Complex'. *Journal of the American Psychoanalytic Association 27*: 751–775.

Luborsky, L. (1967). 'Momentary forgetting during psychotherapy and psychoanalysis: A theory and research method'. *Psychological Issues 5*(2–3): 175–217.

Luborsky, L. (1973). 'Forgetting and remembering (momentary forgetting) during psychotherapy: A new sample'. *Psychological Issues 8*: 29–55.

Mahler, (1968). *On Human Symbiosis and the Vicissitudes of Individuation*, Vol. 1, *Infantile Psychosis*. New York: International Universities Press.

Malinowski, B. (1927). *Sex and Repression in Savage Society*. London: Routledge & Kegan Paul.

Masling, J. (ed.) (1983). *Empirical Studies of Psychoanalytic Theories*, Vol. 1. Hillsdale, NJ: Analytic Press.

Masling, J. (ed.) (1986). *Empirical Studies of Psychoanalytic Theories*, Vol. 2. Hillsdale, NJ: Analytic Press.

Masling, J. (ed.) (1990). *Empirical Studies of Psychoanalytic Theories*, Vol. 3. Hillsdale, NJ: Analytic Press.

Masling, J. M. and Bornstein, F. R. (eds) (1993). *Empirical Studies of Psychoanalytic Theories*, Vol. 4, *Psychoanalytic Perspectives on Psychopathology*.Washington, DC: American Psychological Association.

(p. 374) Masling, J. M. and Bornstein, F. R. (eds) (1994). *Empirical Studies of Psychoanalytic Theories*, Vol. 5, *Empirical Perspectives on Object Relations Theory*. Washington, DC: American Psychological Association.

Masling, J. M. and Bornstein, F. R. (eds) (1996). *Empirical Studies of Psychoanalytic Theories*, Vol. 6, *Psychoanalytic Perspectives on Developmental Psychology*. Washington, DC: American Psychological Association.

Masling, J. and Cohen, I. S. (1987). 'Psychotherapy, clinical evidence, and the self-fulfilling prophecy'. *Psychoanalytic Psychology 4*: 65–80.

Masserman, J. H. (1944). *Behavior and Neurosis*. Chicago, IL: University of Chicago Press.

Masserman, J. H. (1950). 'Experimental neurosis'. *Scientific American 182*: 38–43.

Meehl, P. E. (1973). 'Why I do not attend case conferences'. In *Psychodiagnosis: Selected Papers* (pp. 225–302). Minneapolis, MN: University of Minnesota Press.

Meehl, P. E. (1984). 'Subjectivity in psychoanalytic inference: The nagging persistence of Whilhelm Fleiss Achensee question'. In J. Earman (ed.), *Minnesota Studies in the Philosophy of Science*, Vol. 10, *Testing Scientific Theories* (pp. 349–412). Minneapolis, MN: University of Minnesota Press.

Meehl, P. E. (1993). 'If Freud could define psychoanalysis, why can't ABPP do it? *Psychoanalysis and Contemporary Thought 16*: 299–326.

Meehl, P. E. (1995). 'Commentaries'. *Journal of the American Psychoanalytic Association 43*: 1015–1023.

Mischel, W. (1961). 'Father absence and delay of gratification: Cross-cultural comparison'. *Journal of Abnormal and Social Psychology 63*: 116–124.

Mitchell, S. A. (1998). 'The analyst's knowledge and authority'. *Psychoanalytic Quarterly 67*: 1–31.

Morehead, D. (1999). 'Oedipus, Darwin, and Freud: One big, happy family? *Psychoanalytic Quarterly 68*(3): 347–375.

O'Shaughnessy, E. (2005). 'Whose Bion? *The International Journal of Psychoanalysis 86*: 1523–1528.

Peterfreund, E. (1983). *The Process of Psychoanalytic Therapy: Models and Strategy*. Hillsdale, NJ: Analytic Press (Erlbaum).

Popper, K. R. (2003/1965). *Conjectures and Refutations: The Growth of Scientific Knowledge*. London: Routledge.

Portuges, S. H. (1994). 'Knowledge knots: Problems of psychoanalytic research in clinical psychoanalysis'. Paper presented at the Annual Spring meeting of the Division of Psychoanalysis of the American Psychological Association. Washington, DC, 16 April 1994.

Pötzl, O. (1917). 'The relationship between experimentally induced dream images and indirect vision'. *Psychological Issues 2*: 41–120.

Reichenbach, H. (1938/1951). *Experience and Prediction*. Chicago, IL: University of Chicago Press.

Remnick, D. (2016). 'An American tragedy', **http://www.newyorker.com/news-desk/an-american-tragedy-donald-trump**.

Robins, L. N. and McEvoy, L. (1966). 'Conduct problems as predictors of substance abuse'. In L. Robins and M. Rutter (eds), *Deviant Children Grown Up* (pp. 182–204). Baltimore, MD: Williams and Wilkins.

Rosegrant, J. (2012). 'Why Bion? Why Jung? For that matter, why Freud? *Journal of the American Psychoanalytic Association 60*(4): 721–745.

(p. 375) Rosenblatt, A. (1989). 'Reinspecting the foundations of psychoanalysis: A rejoinder to Adolf Grunbaum'. *Psychoanalysis and Contemporary Thought 12*: 73–96.

Rosenzweig, S. (1934). 'An experimental study of memory in relation to the theory of repression'. *British Journal of Psychology 24*3: 247–265.

Ryum, T., Stiles, T. C., Svartberg, M., and McCullough, L. (2010). 'The role of transference work: The therapeutic alliance and their interaction in reducing interpersonal problems among psychotherapy patients with cluster C personality disorders'. *Psychotherapy: Theory, Research, Practice, Training 47*: 442–453.

Sachs, D. (1989). 'In fairness to Freud: A critical notice of the *Foundations of Psychoanalysis*'. *The Philosophical Review 98*: 349–378.

Schut, A. J., Casonguay, L. G., Flanagan, K. M., Yamasaki, A. S., Barber, J. P, Bedics, J. D., and Smith, T. L. (2005). 'Therapist interpretation, patient–therapist interpersonal process, and outcome in psychodynamic psychotherapy for avoidant personality disorder'. *Psychotherapy: Theory, Research, Practice, Training 42*(4): 494–511.

Sears, R. R. (1943). 'Survey of objective studies of psychoanalytic concepts'. *Social Science Research Council Bulletin* (No. 51). New York: Social Science Research Council.

Sears, R. R. (1944). 'Experimental analysis of psychoanalytic phenomena'. In J. M. Hunt (ed.), *Personality and Other Behavior Disorders* (pp. 306–332). Oxford: Ronald Press.

Silberschatz, G. (2005). 'The control-mastery theory'. In G. Silberschatz (ed.), *Transformative Relationships: The Control-Mastery Theory of Psychotherapy* (pp. 3–30). New York: Routledge.

Spence, D. P. (1994). *The Rhetorical Voice of Psychoanalysis*: *Displacement of Evidence by Theory*. Cambridge, MA: Harvard University Press.

Taylor, E. J. and Bagby, R. M. (2013). 'Psychoanalytic and empirical research: The example of alexithymia'. *Journal of the American Psychoanalytic Association 61*: 99–133.

Weinberger, D. A. (1990). 'The construct validity of the repressive coping style'. In J. L. Singer (ed.), *Repression and Dissociation* (pp. 337–386). Chicago, IL: University of Chicago Press.

Weinberger, D. A., Schwartz, G. E., and Davidson, R. J. (1979). 'Low-anxious, high-anxious, and repressive coping styles: Psychometric patterns and behavioral and physiological responses to stress'. *Journal of Abnormal Psychology 88*: 369–380.

Weiss, J. and Sampson, H. (1986). *The Psychoanalytic Process: Theory, Clinical Observation, and Empirical Research*. New York/London: Guilford Press.

Whiting, J. W. M. (1960). '*Totem and taboo*: A re-evaluation'. In J. Masserman (ed.), *Psychoanalysis and Human Values* (pp. 150–154). New York: Grune and Stratton.

Whiting, J. W. M. and Child, I. L. (1953). *Child Training and Personality: A Cross-Cultural Study*. New Haven, CT: Yale University Press.

Winnicott, D. W. (1953). 'Transitional objects and transitional phenomena.' *The International Journal of Psychoanalysis 34*: 89–97.

Winnicott, D. W. (1958). 'The capacity to be alone'. *International Journal of Psychoanalysis 39*: 416–420.

Winnicott, D. W. (1960). 'The theory of the parent–infant relationship'. *The International Journal of Psychoanalysis 41*: 585–595. (p. 375) (p. 376)

Notes:

([1]) As Meehl (1995) has observed, there probably is no presumed therapeutic intervention in the history of humankind, including bloodletting, purging, and insulin coma therapy, that has not been associated with personal testimony and anecdotes regarding its beneficial effects.

([2]) I would add that accountability is a moral as well as an empirical issue. That is, the clinician has the moral responsibility to be open to the issue of accountability and to avail him- or herself of whatever evidence is available regarding therapeutic outcome and the processes linked to outcome. This is not to say that there cannot be open and informed debates about such issues as criteria for outcome and quality of evidence, including its ecological validity.

([3]) Ironically, despite a widespread hostility towards research, meetings at the American Psychoanalytic Association are labelled 'scientific' meetings.

([4]) In addition to providing a poignant account of the doubts and struggles experienced in psychoanalytic training by an open-minded psychoanalyst, Portuges also contributes an excellent discussion of the scientific status of psychoanalysis as well as of the relationship between research and clinical practice.

([5]) [Eds: For a discussion of Lacan's redefinition of the Oedipus complex, see Boothby (this volume).]

([6]) As one of the editors (RG) notes, one could hypothesize that an exacerbation of A. G.'s symptoms in reaction to 'doing things his father used to do' may reflect an unconscious anxiety-ridden desire and fantasy to replace father. That is a plausible clinical hypothesis not necessarily contradicting, but perhaps complementing, my formulation. However, it is a hypothesis that further clinical evidence would either support or fail to support.

Morris N. Eagle

California Lutheran University, USA

Subscriber: UC - Los Angeles (UCLA); date: 17 March 2019

Oxford Handbooks Online

Abstract and Keywords

Core concepts of psychoanalysis admit integration with recent work in neuroscience and developmental psychology. Karl Friston's free energy neuroscience embodies analogues of Freud's primary and secondary processes, and despite conceptual differences assigns the same causal role as Freud to the minimization of free energy. This supports parallel psychoanalytic and neuroscientific accounts of development, dreams, and symptoms, with the latter expressible as a complexity theory of dreaming and mental disorder. Overall the generative model that performs predictive processing discharges the functions—including the regulation of waking and dreaming consciousness—that Freud assigned the ego, while the prototype emotions systems delineated by Jaak Panksepp and other affective neuroscientists discharge those of the drives or id. Together with the emerging role of REM dreaming in memory consolidation, this provides neuroscientific underpinning for psychoanalytic conceptions of the role of memory, emotion, and phantasy in dreams, as illustrated by Freud's dream of Irma's Injection.

Keywords: predictive processing, free energy, ego, drives, complexity, conflict, dreaming, memory consolidation, phantasy

Introduction: Integrating the Sciences of the Mind

Recent developments have made it reasonable to think that some main claims of psychoanalysis can be relatively directly integrated with the framework of contemporary computational and affective neuroscience. Other approaches in psychology, including attachment theory, are also seeking integration with neuroscience; and both neuroscience

and the study of mental disorder are themselves undergoing integration with genetics and Darwinian evolutionary biology. Here we will focus mainly on psychoanalysis and attachment, leaving evolution for consideration elsewhere.[1]

Probabilistic Regulatory Modelling

The integration advanced here is based on the notion of *probabilistic regulatory modelling* associated with the *free energy principle* recently espoused by Karl Friston and his colleagues (see, e.g. Friston 2010).[2] As accords with Conant and Ashby's (1970) proof that a good regulator of any system must model that system, we can understand Darwinian *adaptation* as shaping phenotypic traits of organisms (and therewith the genes that sustain them) to model aspects of the environment so as to interact with them in ways that promote reproductive success. Taken overall these mappings constitute organisms themselves as dynamic models of the environments with which they interact.

In the nineteenth century—and as a scientific version of Kant's claim that consciousness arose from a conceptual synthesis of the manifold of sensory intuition—Helmholtz proposed that the brain naturally conducted such modelling. He argued that the brain accomplished synthesis by embodying a *generative model* of the causes of the data at its sensory receptors—that is, a model of the causes of these sensory data that could predict the impingements and recreate their effects. (Hence, as Helmholtz remarked, objects were 'imagined in the field of vision' as they would have to be 'to produce the same impression on the nervous mechanism' (1865/1924: 2)).

The hypotheses of this model, as in Kant's account of synthesis, included our concepts of the causes and objects of everyday perceptual experience. In Helmholtz's version our conceptual apparatus *unified* and *predicted* the statistical patterns of the data impinging at the receptors, and so explained these data *for the subject*, by *re-representing* the patterns as *conscious perceptual experiences of their causes*.[3] In this Helmholtz used Kant's philosophy to provide an outline account of the working of the brain in generating consciousness, and one which entailed that the basic sensory data were *not* conscious experiences themselves, but rather the contacts at the neural receptors whose patterns the experiences serve to represent and predict.[4]

Freud's Free Energy Neuroscience

Freud framed many of the basic claims of psychoanalysis in the autumn of 1895, shortly after he discovered the use of free association in analysing his own and his patients' dreams. After this both he and his patients sought self-understanding by following and

describing the rapidly changing contents of their own conscious states, in as much detail as possible and without omission or censorship.

This proved a uniquely informative mode of self-disclosure, which even now remains without parallel in any other discipline. An individual's associations made it possible to consider a whole range of what would emerge as motivated behaviours (verbal utterances, non-verbal actions, thoughts, dreams, and symptoms) together and in relation to one another and to the motives that gave rise to them.[5] Together with his own self-analysis this enabled Freud to learn as much about his patients' states of mind as they were able to put into words, and to extend this by framing and testing hypotheses about what their (and his own) associations indicated they could *not* put into words.

Accordingly he rapidly discovered that 'the pathological mechanisms which are revealed in the most careful analysis in the psychoneuroses bear the greatest similarity to dream-processes' (Freud 1895/1950: 336).[6] He began to shift his investigative and therapeutic focus away from what he and Breuer had taken to be veridical but repressed memory, and towards the fundamental and memory-distorting role of imagination and phantasyy.

Freud's first love in research had been the study of the nervous system, in which he had shown unusual distinction.[7] Hence as his new data prompted new hypotheses, he initially tried to formulate them in neuroscientific terms. Within a few weeks he had framed a prescient account of the brain as operating to minimize a form of *free energy* that was specific to the nervous system.

Freud's idea must have originated as a version of Helmholtz's conception of free energy as *the energy in a system that was available for conversion into work*. All physical engines (animal, vegetable, or mineral) operate in accord with the laws of thermodynamics in converting Helmholtz free energy into work. In organic metabolism this conversion is not straightforward, since it involves concurrent processes that are bidirectional, that is, both energy-storing and energy-releasing. But in applying Helmholtz's notion to the brain and mind, Freud adapted it in ways that we can now see to coincide with the distinct notion of *variational free energy*, as this is used in contemporary neuroscience.[8]

Variational Free Energy (FE)

There is a fundamental conceptual problem in applying Helmholtz's account of modelling to the brain. Predicting sensory impingements and recognizing their causes requires computing, for all relevant causes and impingements, the conditional probabilities of *impingements given causes* that enable the prediction of impingements, and the (inverse) probabilities of *causes given impingements* that enable the recognition of causes. According to the Bayesian mathematics of probability, these calculations are intractable: they have no finitary (closed-form) solution. Thus although Helmholtz had described a

process of modelling that the brain seemed to be performing, the task seemed impossible for the brain to perform.

Geoffrey Hinton and his colleagues solved this problem in a series of papers on the *Helmholtz machine* (Dayan et al. 1995) that marked the transformation of connectionist machine learning into the computational neuroscience of the Bayesian brain. Describing the human sensory system as 'a statistical inference engine whose function is to infer the probable causes of sensory input' (1995: 1), they produced a proof of principle as to how the requisite inferences might be performed. This turned on the notion of *variational free energy* (hereafter FE) that they used in computational modelling.

It could be proved that minimizing the FE of a Bayesian probabilistic model of the causes of some data minimized the divergence in probabilities between those the model assigned to the causes and those of the causes themselves: at zero free energy the probability distributions would be the same. Also, however, the parameters (hypotheses) that determined the FE of a model could be adjusted from within that model (as happens in the brain in perceptual learning) in computationally tractable ways. As Hinton realized, it followed that a brain that embodied such a model as Helmholtz had envisaged could ensure that this model enjoyed maximal evidence-based conformity with the causes it represented by continually adjusting its parameters to minimize FE.

The FE of a model is equal to its *computational complexity* minus its *predictive accuracy*. Both are measured with respect to samples of the data the model adjusts itself to learn to predict, which we can take as the sensory contacts that occur over a single waking period. Accuracy is a measure of predictive success. If we imagine the data of waking impingement as points on a graph, and the FE-minimizing hypotheses generated by the model as curves through those points, accuracy would be maximized by a curve that passed through every data point. Complexity, by contrast, reflects the *changes* imposed on the model in working to increase accuracy, as measured by the number of parameters the model is compelled to *adjust* and the extent of the adjustment required. The brain's model thus minimizes FE by *maximizing accuracy* and *minimizing complexity*, thereby tending to make itself the *simplest best predictor* of the data at the receptors.

Friston extended this framework from the perceptual systems to the overall regulation of bodily change and movement that perception subserves. Thus it appeared that the brain so animated the body as to minimize the variational FE of its regulatory model not just by improving prediction (as in perceptual learning), but also by *framing and satisfying predictive hypotheses as to how optimally to move and act*.[9] Such a minimization of FE constitutes life-sustaining activity that is subserved by the processes that store and release thermodynamic free energy (Seth and Friston 2017). In this sense the minimization of FE takes the foreground in understanding the overall energic working of brain and mind.[10] In consequence, as Helmholtz had claimed, 'every movement we make' can rightly be regarded as an experiment testing the predictive success of brain and body in modelling and navigating their environments.

For both Freud and Friston the free energy to be minimized enters the nervous system via sensory impingement, the main source being the inescapable flow of 'endogenous [*interoceptive*] stimuli' that reflect 'the peremptory demands of the internal needs' (Freud 1895/1950: 297). For Friston these stimuli predict departure from the continuously recalculated targets (including homeostatic set points and allostatic adjustments) for FE-minimizing activity (Pellouzio et al. 2015). Hence as Freud speaks of such inputs as creating a 'demand for work' to produce 'specific actions' that 'bind' (inhibit) this free energy, so Friston speaks of 'an imperative to minimize prediction error ... through action' (2012: 248) that likewise inhibits the interoceptive sources. And on both accounts such minimization requires the brain to embody a representation or model of the world, including the agent's body (in Freud the 'bodily ego', bounded both internally and externally by the skin), via which the minimizing activities are conducted.

These similarities, as well as others described in what follows, give Freud's conception of free energy the same overall causal profile as Friston's FE. On this basis Carhart-Harris and Friston (2010) sought 'to demonstrate and develop the construct validity of the Freudian concepts', including those of the primary and secondary processes. In what follows, therefore, we will advance their argument by using 'FE' to designate both conceptions of free energy.

FE and Waking and Dreaming Consciousness

Freud hypothesized that both dreams and symptoms functioned to protect the self (or 'ego') from trauma and conflict. They did so, moreover, in the same way: via the fabrication of fictive experiences and beliefs, forms of imagination or phantasy that created an 'alienation from reality' (Freud 1911b: 218ff) that temporarily masked and pacified the conflict or trauma involved. Dreaming, moreover, seemed to present an innate form of consciousness, which might perform a similar function in infancy.

Accordingly Freud hypothesized that the original form of consciousness was *imaginary*, and produced by a *primary process* that was 'in the apparatus from the first' (1900: 603). In infancy this process operated in accord with the *pleasure principle*, buffering the potentially traumatic influx of FE at birth, and paved the way for the *secondary processes*, which operated in accord with the *reality principle*. These enabled the infantile ego to distinguish perceptual experience from wishful imagining, to establish experiential memory, and to learn to perform the 'specific actions' required to inhibit the internal sources of FE. Freud illustrated this by an example in which an infant used the 'wishful cathexis' provided by imagining the breast, to 'experiment' by turning her head so as to secure the nipple (1895: 328–9, 356). This required the infantile ego to moderate (inhibit) the relief provided by phantasy, in order to focus on the *indications of reality* that enabled movement leading to a veridical *experience of satisfaction*.

Thus against the background of the regulation of pleasure and unpleasure by the primary process, the secondary processes introduced the infant—cry by cry, feed by feed, excretion by excretion—to the harsh realities of the world into which she had been born, and also to her own real resources for coping with them. So as the infant gained a greater apprehension of reality (and to the extent that she did so) the secondary processes inhibited and overlaid the primary process in waking life, relegating it to the motor paralysis and sensory attenuation of dreaming.

This benign development, however, remained under threat from frustration, trauma, or conflict. These tended to activate the primary process in waking, and so could create a continued reliance on wishful phantasy that would later show in the alienation from reality (Freud 1911b) characteristic of mental disorder. Hence as Freud later described matters, the phantasies produced by the primary process 'possess *psychical* in contrast to *material* reality; and we gradually learn to understand that *in the world of the neuroses it is psychical reality that is of the decisive kind*' (Freud 1915–7: 368).

The contemporary FE framework provides a remarkably full parallel to these claims. Sleep is a key factor in neurological development, and infants cycle through phases resembling slow wave (SWS) and rapid eye movement (REM) sleep while still in the womb (Kurth et al. 2015). In the last trimester they spend most time in REM, and as Reissland et al. (2011, 2013) observe, their facial expressions indicate both positive (laughter-related) and negative (cry-related) emotions. Accordingly Hobson, Hong, and Friston (2014) take this as a formative period for the brain's generative model. They hypothesize that the brain is 'genetically endowed with an innate virtual reality generator' whose working 'is most clearly revealed in rapid eye movement sleep dreaming'.

We are thus 'born with a virtual reality model' of what we will later discover to be the worldly causes of sensory impingement; and with the onset of sensory activity this virtual model is 'entrained by sensory prediction errors', to become 'a generative or predictive model of the world'. In Friston's account this is accomplished by *active inference*, which effects the perceptual learning, establishing of memory, and inception of FE-minimizing action that Freud assigned to the secondary processes.

This account also parallels post-Freudian (e.g. Kleinian) descriptions of the primary process as *innate expectation-generating phantasy*. Thus Segal (1964) describes the infant as 'approach[ing] reality armed, as it were, with expectations formed by his unconscious phantasy'. These expectations are comparable to scientific hypotheses, in the sense that 'by testing them in reality [the infant] gradually learns which are applicable and which modes of his own functioning enable him to deal with reality'. (Among them are the 'preconceptions' of ideally good and malignantly bad breast/mother/ others that the Kleinian tradition regards as the innate basis for thinking about ourselves and others.)

In Friston's terms these expectation-inducing would be the *prior hypotheses* that are altered to accommodate prediction error as the virtual reality generator becomes a 'predictive model of the world'. Alteration of these priors in the first phase of perceptual learning would produce experientially informed *posterior hypotheses* that would then serve as priors for the next phase.

In all this Helmholtz, Segal, Bion, and Friston construe *thinking* in a way that has become familiar in philosophy—as involving tacit or explicit *abduction*, or *inference to the best explanation* (Douven 2011). Helmholtz and Friston take perceptual consciousness to be created by tacit abduction on the part of the brain or generative model; and as conscious agents we continue abductive inference in everyday life and in science. This is the continuity, as Popper remarked, that enables human beings to 'let their theories die in their stead'.[11]

These accounts clearly impose a common framework that links waking, dreaming, development, and mental disorder. On both accounts waking consciousness is underlain by an original imaginary (virtual reality/phantasy) process that later appears mainly in dreaming, and whose operation in waking is constrained by a model or system of representations of the causes of sensory impingement. On both, therefore, a central aspect of development consists in constructing the worldly model whose adherence to reality inhibits and overlays the imaginary process in waking. This, however, entails the possibility originally envisaged by Freud, namely that the development of this realistic orientation may be more or less unsuccessful, leaving the imaginary process to produce the alienation from reality that appears as mental disorder.[12]

Friston's account of FE casts further light on these matters. We noted earlier that the brain's model minimizes FE by maximizing accuracy and minimizing complexity, thereby tending to make itself the simplest best predictor of the data at the receptors. In Friston's account these tasks are roughly divided between waking and sleeping. As a first approximation (omitting the role of imaginative/primary process activity in waking) the brain operates to increase the accuracy of the generative model in waking, and to reduce its complexity in sleep and dreaming.

In this context the most accurate model would be the one best able to predict the currently optimal course of motor response and put it into action. The contrasting importance of complexity emerges if we consider that if an agent's model is too complex —for example if it is working to satisfy imperatives that conflict with one another—such computation may be difficult or impossible. An agent who *needs* a drink, for example, will maximize accuracy by *wanting* and *getting* a drink. But an agent who has numerous conflicting needs may be unable to effect satisfying computation of this kind, and so remain effectively immobile. Hence the requirement, advanced over the lifetime by processing in sleep, of simplifying to predict a single best alternative. As we will see further on, the reduction of complexity in dreaming apparently takes place via the

imaginary attainment of accuracy that Freud described as wish fulfilment (see Pataki, this volume).

We return to dreaming in 'The Complexity Theory of Dreaming and Mental Disorder'. First, however, we must take up the relation of complexity to emotion and conflict.

Emotion Systems

In regulating waking and dreaming consciousness in such a way that experience, thought, and action harmonize to minimize FE, the generative model discharges the tasks that Freud assigned to the ego. In this, however, it is important to see that the brain as modeller or ego is not just a *functional part* of the personal conscious self, but is rather the *overall regulatory governor* of the organisms we are. Among its activities are the production of the conscious experiences in which we regard ourselves as persons who think, feel, choose, and act, as well as the dreams in which we imagine ourselves doing so.

The locus of this ego- or self-creating regulatory activity is the cerebral cortex; and the *subcortical prototype emotion systems*[13] that the cortex develops to regulate over the first year of life play the role that Freud assigned to the drives or id. By contrast with the cortex, the subcortical generators of emotion are fully functional at birth, and their basic architecture in the brainstem is common to all mammals. Watt and Panksepp (2009: 92–3) describe these systems as 'sitting [in the brainstem] over homeostasis proper', where they give rise to attachment, which becomes 'the massive regulatory-linchpin system of the human brain', exercising '*primary* influence' over the prototype systems below.[14]

As described in Alcaro and Panksepp (2011) and Alcaro et al. (2017), these basic systems divide into two overarching and conflicting groups:

1. Those characterized as 'positive/appetitive' or 'Approach/Reward'. As indicated by Panksepp's capitalized functional descriptions, these include SEEKING, PLAY, LUST, and CARE.

2. Those characterized as 'negative/aversive' or 'Avoidance/Conflict'. These include RAGE and FEAR, as well as the attachment-regulating system that Panksepp describes in terms of PANIC, SEPARATION DISTRESS, and GRIEF (hereafter PSG). This latter system generates the punishing internal experiences that ensure the proximity of reproductive cooperators such as parents and offspring.

The systems in each group are co-activating (RAGE and FEAR activate one another, and PSG activates both, and the objects of SEEKING include cooperators in CARE, PLAY, and LUST). By contrast the groups inhibit one another, acting as parts of an overall structure of emotional opponent-processing (Craig 2015); and they generate conflicting dispositions in physiology and behaviour. They are thus sources both of the opposition between

unpleasure and pleasure stressed by Freud, and of the emotional conflicts that he studied and aimed to mitigate—and that we can now see as contribute to insecure (and particularly disorganized) attachment.

These systems coordinate in reducing FE, but via a dialectic that renders them liable to precipitate emotional conflict. This begins with birth, when the negative/aversive systems are activated by the prediction errors that signal impending violations of basic homeostatic expectations previously met from within the mother's body. This challenge to life-sustaining equilibrium is expressed in what Freud regarded as the infant's 'scream'. This communicative action is often the infant's first, and produces a resonating activation in the aversive systems of the mother (Parsons et al. 2014), galvanizing her response. Sensitive response (Howe 2011) coordinates mother and infant in replacing aversive activations with appetitive and rewarding SEEKING, approach, and the consummatory and life-sustaining activity of nursing. This promotes synchronization and bonding between them (Atzila et al. 2017), paving the way for the other approach/reward systems —LUST, CARE, and PLAY—to participate in deepening their emotional relationship.[15]

In this dialectic the infant's and mother's generative models operate in the same way. Equilibrium-threatening increases in FE activate the avoidance/punishment systems, and the model responds with motor activity (an optimal trajectory) that engages the approach/reward systems to minimize it. In infancy this pattern coordinates mother and infant, but the pattern itself is paradigmatic for the operation of the generative model, whether operating realistically or in virtual reality. It develops together with attachment over the first year, as the cortex extends its governance of the emotion systems through phases of axonal growth, synaptic proliferation, and experience-dependent neural pruning. By the beginning of the second year the brain has roughly doubled in size, and reliable patterns in attachment relationships can be discerned in the 'Strange Situation' procedure, to be discussed in the section 'Conflict and Complexity in Disorganized Attachment' further on (see also Howe 2011).

This marks the stabilizing 'regulatory linchpin' of the infant's establishing of 'internal working models' for emotionally significant relationships with the mother and/or other carers. This modelling discharges what Freud described as 'one of the earliest and most important functions of the mental apparatus' namely 'to bind the instinctual impulses which impinge on it, to replace the primary process prevailing in them by the secondary process and convert their freely mobile cathectic energy into a mainly quiescent (tonic) cathexis'' And as Freud also remarks, and we will see later, 'failure to effect this binding would provoke a disturbance analogous to a traumatic neurosis ...' (Freud 1920: 33, 34).

Trauma, Conflict, and Complexity

Infancy combines early neurological development with experiential learning, and is therefore uniquely formative for the brain. For this reason early adverse or traumatic experience can have a lifelong impact on mental and physical health (Anda et al. 2006) that is mediated by changes in the working of the brain (McCrory et al. 2017). Early activation of the aversive emotional systems seems a likely cause of what Freud described as 'the first great anxiety-state' associated with the trauma of birth, and later 'the infantile anxiety of longing—the anxiety due to separation from the protecting mother' (Freud 1923: 58).

The early focus of the conflicting emotion systems on the mother accords with psychoanalytic claims that experience in early infancy alternates between prototype-forming representations of very bad and very good objects, whose originals are the parts of the mother's body with which the infant has significant interactions—her breast, her holding presence, her face, eyes, etc. Recent research on memory suggests that the infant's episodic (autobiographical) memory system, centred in the hippocampus and amygdala, is 'highly engaged in the processing of infantile experience and the storing of latent memories' (Alberini and Travaglia 2017; see also Travaglia et al. 2016). These latent memories, moreover, seem comparable to the early 'memories in feeling' described by Melanie Klein (1952: 234). For although they do not admit explicit recall, the emotions linked with them are nonetheless rearoused when similar experiences occur at later times. Despite their latency, such memories can play an important role in experience-driven maturation and development.

This is important for our purposes, for Friston's conception of complexity is internally connected with both *trauma* and *conflict*. Complexity measures *the burden of adjustment* that experience places on the generative model; and since the primary task of the model is to calculate an optimal motor trajectory, as driven by the potentially conflicting prototype systems, this is effectively the *burden of emotional adjustment* (cf. the notion of *affective load* in Levin and Nielsen 2009). Conflict naturally creates complexity, since it involves the activation of *inconsistent* (conflicting) patterns in behaviour, one or both of which perforce require complexity-reducing adjustment. And experiences are traumatic when—as in post-traumatic stress disorder (Enlow et al. 2014) or borderline personality disorder (Mosquera et al. 2014)—they impose requirements for emotional adjustment that the ego/generative model cannot fully meet. The overarching role of the potentially conflicting prototype systems in human motivation thus ensures that for *our* generative models complexity is mainly *emotional* complexity, even when (as, say, in obsessive compulsive disorder or OCD) it shows in ostensibly cognitive forms.

Conflict and Development

We have seen that the early working of the emotion systems can engage mother and infant in a conflictual relationship that predates the infant's understanding. The activation of conflicting emotions becomes emotional conflict *for the infant as subject and agent* with the recognition of self and other as distinct enduring individuals. Experiments suggest that this development—and with it our conception of ourselves as unique persons bounded by physical bodies that move in space and last through time—starts to take hold in the fourth month, as Melanie Klein suggested in her account of the depressive position.[16]

In the prior, pre-objectual phase (Kumin 1996) infants seem to conceive the world as containing a numerically indeterminate multitude of *episodic proto-objects*, comparable to the 'shifting and unsubstantial "tableau"' in terms of which Piaget and Inhelder (1969) described early experience.[17] Since these proto-objects are derived from perception of parts or aspects of the objects perceived, these are also 'part objects' which are anatomically and psychologically incomplete;[18] and since the infant's conception does not yet encompass their actual patterns of spatio-temporal behaviour, they are particularly liable to imaginary transformation in phantasy.

During this phase the infant is also learning the boundaries of the bodily self via sight, touch, and movement, as well as a range of feelings generated via interoception (Fotoupoulou and Tsakaris 2017).[19] So it is *also* indeterminate for the infant whether these episodically encountered proto-objects are mental or physical, or inside or outside the bodily container of the mind or self, as this is bounded by the skin. In consequence the infant can imagine taking the early proto-objects into itself with milk or food; as good or bad causes of feeling inside the self; as expelled from the self, or put into others, with urine, gas, etc. and so on. (Such phantasies populate dreams, and often re-emerge in mental disorder.)

As the generative model extends its spatio-temporal scope the infant increasingly conceives persons as unique and persisting, and accordingly merges proto-objectual representations to represent good and bad aspects of single complex individuals. This requires reorienting the aversive systems from their original focus on the mother, apparently effected by a coordination of identification, projection, and attachment. The realization that others are unique and irreplaceable spurs identification with their valued and caring aspects, and fosters the infant's capacities for guilt and concern, ameliorating the ingroup-facing aspects of the superego. At the same time aspects that cannot be integrated in this way—particularly those derived from early parent–offspring and sibling rivalry—are split off and projected *outside* the circle of familiar kith and kin.

These developments also establish the family as a basic ingroup, preparing the infant for ingroup cooperation for outgroup conflict. Some early signs of this appear at seven to eight months, when infants start to show both angry protest at separation from their

mothers and also fear of strangers (Schaffer 1971). The separation protests show appreciation of the mother as unique and irreplaceable, as well the refocusing of the emotion systems that this entails. RAGE at the mother now operates in harmony with PSG and FEAR (of strangers) to maintain proximity and attachment within the family. And the projections that originally created the indeterminately numerous bad proto-objects of early infancy are relocated in the alien stranger, and so can now operate to arouse derivatives of RAGE and FEAR in ingroup cooperation for outgroup conflict.[20]

Conflict and Complexity in Disorganized Attachment

We have been taking the early binding (successful inhibition) of free energy hypothesized by Freud to be accomplished by establishing the emotional bonds of attachment. But as noted in the section 'Emotion Systems', Freud also claimed that 'failure to bind' [successfully inhibit] the FE of the drives 'would provoke a disturbance analogous to a traumatic neurosis ...' (1920: 64). We can see such a failure, and learn something of its causes, by considering disorganized attachment, as shown in the Strange Situation.

This is a roughly twenty-minute procedure involving eight episodes, numbered (1) to (8); and we can regard the assignment of categories of attachment by this procedure as measuring the functional integration of the rewarding and aversive prototype systems. In (1) the mother (or other carer) and infant enter the room, and in (2) the infant has the opportunity to explore and play with the toys which are there. In (3) a stranger enters the room, and after a short interval attempts to play with the infant. In (4) the mother goes out of the room, leaving the infant with the stranger, who tries, if play has begun, to continue it. Then in (5), the first reunion, the mother returns and the stranger leaves while the mother settles the infant, providing opportunity to return to exploration and play. In (6) the mother again goes out, this time leaving the child entirely alone, and in (7) the stranger enters and again attempts to play with the child. Finally, in (8), the second reunion, the mother again returns to settle the child, who may (or may not) go back to exploration and play.

Thus while episodes (2) and (5) and (8) are such as to activate the rewarding systems that Panksepp identifies as SEEKING and PLAY in the context of maternal CARE, (3) and (4) and (6) and (7) are such as to activate aversive FEAR, PSG, and RAGE, with the latter directed at the caring mother for her part leaving the infant alone and/or at the mercy of the stranger. These conflicting activations, moreover, are such as to produce *regression*, since they repeat those that marked the infant's appreciation of the mother as unique and irreplaceable at seven to eight months. Roughly, infants who are classed as secure at twelve months show less conflict between the rewarding and aversive systems, and

resolve their conflicts more readily, than those who are classed as insecure, that is, as avoidant, ambivalent-resistant, or disorganized.

Infants classed as disorganized show the most striking and apparently irresolvable conflicts. Here are two examples, from a girl and a boy aged eighteen months, as described in Solomon and George (1999: 131).

> In the second reunion Kate approached her mother with her arms outstretched … when she was about two feet away from making contact, she moved her arms to the side and abruptly circled away from her mother like a banking airplane. As she moved away she had a blank, dazed expression on her face.

> In the first reunion, Sam approached his mother with his eyes cast down. When he was about two feet away he looked up at her, rising suddenly and making gasping noises with his breath as he did so. He quickly looked down again, bared his teeth in a half-grimace/half-smile and turned away. Hunching his shoulders and holding his arms and legs stiffly, he tiptoed to the other side of the room. He sat motionless in the chair for 30s, grasping the armrests and staring straight ahead with a dazed expression.

This is an example of the kind of complexity—and the overall trade-off between complexity and accuracy—generated by emotional conflict. The successive stages of the Strange Situation have aroused such disparate and conflicting images of these toddlers' relationships with their mother that their egos/generative models are now unable to determine whether the more accurate response would be to approach or avoid her. On the present account this is because their online working memories are activating disparate and conflicting pre-objectual representations of self and other—representations that their egos/generative models have not yet integrated into single coherent representations of themselves in relations to others as single enduring objects. Hence these representations are overly complex in a literal sense. If we compare the generative model to the guidance system of a self-driving vehicle, then these toddlers would be attempting to compute a route on the basis of disparate and conflicting maps (internal working models) of the territory they are seeking to navigate.

Disorganized infants characteristically achieve a particular form of coherence by three to five years, when they settle into patterns of *coercion* and *control* in relation to their objects. Thus here is Kate at forty-two months playing with a little boy named Trey:

> Pretending to bake a cake in a toy oven, Kate said in a very loud, bossy tone, 'You can't have cake now! Go away: you can't have cake 'till I call you!' She then ordered Trey to get some dishes for her. She frowned when she saw what he brought and scolded him, shaking her finger with her hand on her hips. Trey pretended to eat … Kate yelled 'No! It's not done!' … She ordered him to go away … when he didn't move she pushed him roughly, and he fell to the floor …

(Solomon and George 1999: 150).

And here is Sam at the same age with a friendly little girl called Jenny, who tries to interest him in playing with her doll.

> At first he ignores her overtures but finally he takes the doll she offers … [and] alternates between brushing its hair tenderly and smashing it against the floor. Jenny tries to integrate Sam's behavior into her pretend play … she pretends the doll is saying in a squeaky voice 'Ow, don't do that!' … Jenny tries again, making her doll say 'I'm going on the bus'. Sam grabbed the toy bus before Jenny could put her doll in it, crammed his doll in, and drove the bus away. Jenny gave up trying to play with Sam

(Solomon and George 1999: 153).

Here the kind of conflict Sam showed towards his mother at eighteen months seems echoed in his attitude towards Jenny's doll, which he alternatively brushes tenderly and smashes against the floor.

Finally we can see what seem to be pre-objective precursors of this kind of complexity in the microanalyses of videotaped face-to-face interactions between mothers and infants described in Beebe and Lachman (2014). These indicate that disorganization in the Strange Situation at twelve-plus months can be predicted from the fourth month—that is, near the inception of the shift to objectual representation the defines the depressive position—by short (2.5 second) episodes in which the mothers involved *fail to recognize or coordinate empathically with their infants' expressions of distress*. Instead the mothers involved respond to their infants' distress with exaggerated surprise or smiling, or again, as distress increases, with facial freezing and/or looking away. Meanwhile the infants themselves—as in the later disorganized patterns discussed previously—respond with *rapidly conflicting expressions of emotion (*e.g. whimpering and smiling within the same second or so); and they grow more distraught as no response is forthcoming.

These are episodes in which the mothers concerned are *not* meeting activations of their infants' aversive systems in the paradigmatic way discussed previously, that is by responses that generate FE-minimizing approach/reward. Rather they seem comparable to what Bion (1962) and Segal (1975) describe as maternal failure to respond in a *containing* way to infants' projections of distress, leaving infants unable to incorporate a capacity for conflict-modifying containment themselves. Emotional misrecognition of this kind can occur in otherwise caring relationships, particularly, it seems, if the carers experienced something comparable in infancy. Still it can prove traumatic, in the sense of contributing to the kind of complexity found in disorganized attachment, and also, say, in borderline personality disorder (BPD), in which histories of trauma and/or disorganized attachment are common (Mosquera et al. 2014).

The Complexity Theory of Dreaming and Mental Disorder

In the *Project* Freud observed that the processes which produced the 'alienation from reality' of mental disorder had 'the greatest similarity' to those that produce dreams. Hobson, Hong, and Friston (2014) argue that the 'virtual reality' dream processes, together with REM (rapid eye movement) sleep and SWS (slow wave sleep), serve to minimize FE by *reducing the complexity* of the generative model. We have seen that this complexity encompasses emotional conflict and trauma; so the virtual reality generator would be doing the same thing—*generating virtual reality to reduce the complexity introduced by trauma and conflict*—by producing the 'alienation from reality' (Freud 1911b: 218ff) that Freud and many others have found characteristic of mental disorder.

We can thus integrate Freud's and Friston's accounts of dreams and symptoms by advancing a *complexity theory* of dreaming and mental disorder. This is the hypothesis that the conflicts and traumas that Freud thought mitigated by recourse to yphantasy/ virtual reality in both dreaming and mental disorder can be seen as forms of the neurocomputational complexity illustrated by disorganized attachment (Hopkins 2016). The claim is thus that mental disorder is the product of such complexity, together with the mechanisms (which apparently include synaptic pruning) which have evolved to reduce it. As well as incorporating Freud's original hypotheses, this account extends them, by relating waking disorders to aberrant complexity reduction in SWS, REM, and dreaming. As described in Hopkins (2016) this is borne out by research on dreaming in depression as well as by work in schizophrenia and a number of other disorders.[21]

How do sleep and dreaming reduce the complexity of the model, and so prepare it to be a more efficient (simpler) predictor of sensory impingement in the next waking period? We can see this by considering the processes of *consolidation* and *reconsolidation* that establish and update the cortically stored *long-term memories* that inform the model. The 'engram complexes' (Tonegawa et al. 2015a, 2015b) that realize memory are networks of synaptically connected neurons. As Freud anticipated, experiential learning systematically alters the patterns and strengths of synaptic connections (together with other aspects of the neurons that regulate their activity), thus determining the nature and contents of the memories they encode. Those that embody a particular memory are distributed over a number of regions in the brain, which may be connected synaptically and/or coordinated by oscillatory patterns (waves) of various kinds (Boccio et al. 2017).

The working of engram complexes involves two roughly dissociable components: the *contextual content* of the memory (who, what, when, where, etc.), as contributed by encoding activity in the hippocampus; and the *emotional significance* or *valence* of the memory, contributed by encoding activity in the amygdala, as related to the hypothalamus and other subcortical structures. The complexity of what is learned is reflected at the synaptic level in individual neurons. Thus Tonegawa et al. (2015a: 94)

report that conditioning, eliminating, and reconditioning fear memories creates opposing patterns of growth and elimination of spines less than 2 millionths of a metre apart on the same dendrites. This suggests that learning creates dynamic opposition, and hence complexity, even at the level of individual synapses—of which there are trillions in the brain.

Accordingly the creation of learning-embodying synaptic connections puts a physiological load on the neurons involved, and one that increases with complexity. Hence what is called *synaptic homeostasis* (Tononi and Cirelli 2014) requires that these connections be *renormalized* (rescaled to levels optimal for the next waking period) in sleep. Friston (2010) takes this as an important part of the reduction of complexity, and there is now strong evidence that SWS (Tononi and Cirelli 2014) and REM (Li et al. 2017) reduce many synaptic connections while maintaining and strengthen others during sleep. According to recent 'active systems' accounts (Dudai et al. 2015; Feld and Born 2017; Rasch and Born 2013) renormalization is implemented in SWS and REM as part of the overall consolidation and reconsolidation of memory in sleep.

In waking the hippocampus acts in concert with the amygdala as 'the search engine of the brain' (Buzsaki 2011), continually constructing the online working memory that enables 'navigation in *cognitive* [as well as physical] *space*'. This includes 'the computation of relationships among perceived, conceived or imagined items', which depends on memory. In addition the hippocampus rapidly encodes the details of waking experience, so that in SWS it can transfer memories from the day to cortical engram complexes for consolidation as long-term memory. This transfer *reactivates* the complexes, rendering them plastic for integration with the new memories being transferred, and rearouses the associated emotions, so that SWS gives way to REM and dreaming.

In dreaming as in waking, the generative model operates to minimize expected FE, and via activation of the SEEKING system. In waking expected FE arises from sensory impingement and the arousals of memory and emotion engaged in increasing accuracy (e.g. by successful satisfaction of need and desire) to reduce it. In dreaming, by contrast, sensory impingement is curtailed, so that the FE sources for complexity reduction are those of memory and emotion in the process of consolidation dreaming.

For this reason dreaming characteristically involves a pattern in virtual reality that mirrors that in waking, but in relation to the arousal of memory rather than sensory experience. The reactivation of emotion-laden (conflictual or traumatic) memories arouses negative/aversive emotions, prompting the *imaginary* prediction/execution of the FE-minimizing appetitive/rewarding trajectories that Freud described as wish fulfilments. Thus for example Freud's dream of Irma's Injection (1900: 106ff),[22] as we will consider in more detail, begins with Irma complaining to Freud of suffering that his psychotherapy has failed to relieve, but ends with diagnostic triumph on his part.

Accordingly REM dreaming is accompanied by the kind of burst-firing of dopamine neurons that marks waking consummatory reward (Dahan et al. 2007). As this occurs in the conditions of maximal neuroplasticity that enable the reactivated old memories to be integrated with the new, the imaginary attainment of reward in dreaming can reshape the emotional valences of the whole engram complex, rendering it less productive of conflict or trauma, and so emotionally optimized to minimize FE by increasing accuracy during the next waking period. Dreaming thus minimizes complexity by an *imaginary* maximization of accuracy during sleep, which optimizes the brain for actual maximization of accuracy in waking.

As would fit with this, Genzel et al. (2015) argue that REM and dreaming function to modify aversive amygdala-related components of memory.[23] This in turn fits with the long-observed positive effects of the restoration of REM dreaming on depression (Cartwright 2010) as well as with recent experiments on engram complexes (Boccio et al. 2017; Tonegawa et al. 2015a, 2015b) that produce dissociations between their amygdala-related emotional valence and their hippocampus-related contextual contents—and including those in which depression-like activity is reduced by manipulations that arouse reward (Ramirez et al. 2015).

Dreaming thus seems the final phase of both the reduction of neurocomputational complexity and the reconsolidation of memory that occur in sleep (and we would expect these processes to be intertwined). On the present account this reduction/reconsolidation yields *a revised integration of emotion and memory, in which the role of emotionally traumatic or conflictual memories has been mitigated by the imaginary experiences of the dream*.

Dreaming Consciousness and the Reduction of Emotional Complexity

This account integrates the reduction of FE (as complexity) in sleep with observations about the large-scale neurological phenomena of SWS, REM, and synaptic homeostasis, downscaling, and pruning. The role it assigns to the generative model—as *reducing the aversive complexity of emotional memory via the imaginary accuracy of the fictive experiences of dreaming*—explicates the significance of dreams in a way that accords with Freud's as well as many other depth-psychological descriptions. For on this account the significance we assign to dreams derives from their role in *unifying the emotional meaning of the past and the present in preparation for the future*, a task they accomplish by *adjusting emotional significances over the fields of memory in which experiential learning is embodied*.

Psychoanalysis and Neuroscience

We can readily trace this hypothesized adjustment of emotional significance in the examples in Freud's *Interpretation of Dreams* as well as later psychoanalytic writings. For in Freud's work as well as that of other analysts, the analysis of a dream can be divided into three parts, which are related by association, memory, and emotion. These are:

> **i.** The *manifest content* of the dream, the series of fictive conscious experiences that the dreamer reports as making up the dream. Elements of this are related by free association (and memory and emotion) to
>
> **ii.** Memories and feelings *from the day before the dream*, whose contents indicate that they played a causal role in the formation of the dream. These in turn are related by memory, emotion, and association, to
>
> **iii.** Memories from the more remote past, together with deeper and more traumatic and conflicted emotions. The contents of these indicate that they were unconsciously activated together with (ii) the conscious memories and feelings from the day; and these contents can be seen as *transformed* in (i), the manifest content. So Freud describes these as the *latent content* of the dream.

Free association leads from elements of (i) to elements of (ii) and (iii). These connections, and particularly the relations of transformation between (iii) and (i)—what Freud describes as the transformation of latent into manifest content—enable us to gauge the potential impact of the transformations of the manifest dream on the latent memories and emotions that underlie it. We can illustrate this by taking a fragment from Freud's dream of Irma's Injection , and focusing on a single manifest element (which we will follow with the aid of bold type) together with the memories and emotions that Freud associated with it.

(i). From the manifest content of the dream:

> … I at once took [Irma] to one side … to reproach her for not having accepted my 'solution' yet … I took her to the window [to examine her] and looked down her throat … I at once called in Dr. M., and he repeated the examination and confirmed it … [saying] 'There is no doubt it is an infection.' … We were directly aware, too, of the origin of her infection. Not long before, when she was feeling unwell, my friend Otto had given her an injection of …trimethylamin (and I saw before me the formula for this printed in heavy type) … **Injections of that sort ought not to be made so thoughtlessly** …And probably the syringe had not been clean.
>
> (Freud 1900, 107)

This element of the manifest content—Freud's claim in the dream that **Injections of that sort ought not to be made so thoughtlessly**—is clearly associated with memories and feelings, both from the dream day, and from the more remote past.

(ii). Memories and feelings from the day before the dream.

As Freud reports, Irma was a family friend and patient he had been treating for hysteria. As the treatment was ending he had 'proposed a solution to [Irma]' about the nature of her illness 'which she was unwilling to accept', and her sessions were broken off at the summer vacation.

On the day of the dream Freud was visited by Otto, the long-standing friend and colleague whom the dream depicts as giving Irma a toxic injection. Otto had been staying with Irma and her family at their country place, where he had been called away to *give someone an injection*, which event evidently reappeared in the dream in an altered form.

> … I asked him how he had found [Irma], and he answered 'She's better, but not quite well' … Otto's words, or the tone in which he spoke them, annoyed me. I fancied I detected a reproof in them … the same evening I wrote out Irma's case history, with the idea of giving it to Dr M … to justify myself … I had been engaged in writing far into the night …
>
> (Freud 1900: 106).

Here a number of connections between the memories and feelings from the day and the content of the dream are clear. The dream depicts Freud as reproving Irma for not having accepted his 'solution', and explains her continued illness as caused by Otto's having given her a toxic injection. Hence in the dream Freud had transformed his own emotionally aversive daytime experience of feeling himself reproved by Otto on behalf of Irma into one of reproving both Irma and Otto. The **Injections of that sort ought not to be made so thoughtlessly** of Freud's dream was the last and most serious of his self-justificatory retaliations for what he had imagined as Otto's reproof on the day of the dream.

(iii). Memories from the more remote past, together with deeper and more traumatic and conflicted emotions. To **Injections of that sort ought not to be made so thoughtlessly** Freud associated as follows (emphasis again supplied):

> … .this sentence in the dream reminded me once more of **my dead friend** who had so hastily resorted to **cocaine injections** … I noticed too that in accusing Otto of **thoughtlessness in handling chemical substances** I was once more touching upon **the story of the unfortunate Mathilde, which gave grounds for the same accusation against myself** …
>
> (Freud 1900: 117).

These associations go directly to past memories freighted with guilt and shame. As Freud says, they indicate that he felt himself guilty of the same kind of thoughtless handling/ injection of toxic substances of which he accused Otto in his dream, with the addition that in his case injections had been lethal. Moreover, he linked this with his own advocacy of

the medical use of cocaine, which (on the basis of his own experience) he had gone so far as to claim was non-addictive. Thus as he says:

> I had been the first to **recommend** the medical use of **cocaine**, in 1885, and this recommendation had **brought serious reproaches down on me**. The misuse of that drug had hastened the **death of a dear friend of mine**

> (Freud 1900: 111).

The 'dear friend' was Freud's beloved mentor, the talented scientist Ernst Fleichel von Marxow. He had become addicted to morphine, taken to relieve incurable nerve pain from an amputation consequent on dissecting a cadaver. Freud had recommended that he use cocaine as a non-addictive substitute, and his subsequent addiction, decline, and death was a source of grief and guilt. The 'unfortunate Mathilde' whom Freud associated with him was someone about whom Freud felt a comparable guilt. He had given her injections of sulphanol, considered harmless at the time, and she, too, had died as result.

This death, moreover, was linked with a kind of unconscious emotional conflict on Freud's part that he was to explain much later in his career. Using the name 'Mathilde' as a switch-word, he associated the Mathilde he had killed with his own daughter Mathilde, who had also been seriously ill.

> My patient—who succumbed to the poison—had **the same name** as my eldest daughter. It had never occurred to me before, but it struck me now almost like an **act of retribution on the part of destiny**. It was as though the replacement of one person by another was to be continued in another sense: **this Mathilde for that Mathilde, an eye for an eye and a tooth for a tooth**

> (Freud 1900: 111–12).

In considering how his associations had led to these topics, Freud remarked that it seemed he 'had been collecting all the occasions which I could bring up against myself as evidence of lack of medical conscientiousness'. With hindsight we can see that this was the activity of the agentive part of himself that he would later describe as the cruel and moralistic superego, which not only accused him with his most painfully regrettable medical derelictions, but also threatened him with retribution in the form of the death of his own daughter.

This is an example of the turning of aggression against the self to curb aggression within the family that Freud took as one of the main functions of the moralistic superego, and that is particularly evident in the ferocious self-reproaches as well as the disposition to suicide characteristic of introjective depression (Blatt 2004). These memories were thus also a locus of emotional conflict—the conflict between the ego and the moralistic superego in terms of which Freud was later to understand guilt, shame, and suicide.

In psychoanalytic terms we can say that this part of Freud's dream was produced by his ego to relieve guilt and shame generated by his moralistic superego, whose cruelty is shown in the way he feels it to threaten him with the death of his daughter. Similarly in FE terms we can say that the dream was produced by Freud's generative model acting as complexity-reducing virtual reality generator. In this it was operating to reduce the emotional complexity (conflict and trauma, guilt and shame) embodied in Freud's long-term memories, by reconsolidating them under the impact of the imaginary but rewarding and emotionally simplifying experience of innocence and vindication provided by the dream.

Thus in both accounts the mitigation of conflict or reduction of complexity is accomplished by *fictive* experiences, as partly realized by activity of the approach/reward systems, in what we can regard as *instances of self-induced complexity-reducing experiential learning in dreaming consciousness*. And here the agencies involved in the self-induced learning include 'internal objects', or again aspects of the 'internal working models' of self and other, that subserve the working of the ego/generative model.

The FE account thus enables us to describe the role of Freud's memories and associations more fully, and to integrate them with the 'active systems' account of memory consolidation described in the earlier section 'The Complexity Theory of Dreaming and Mental Disorder'. For as already indicated, the memories and feelings in (ii)—those from the dream day—are evidently *those undergoing consolidation in the dream*, that is, those being transferred from the hippocampus to emotionally related engram complexes in the cortex for storage as part of long-term autobiographical memory.

Again, the memories and feelings in (iii) are those that were already resident in the engram complex, whose unconscious activation during the day had set the stage both for Freud's sensitivity to Otto's remark, and for his self-justifying response to it. These are the memories with which the newcomers are undergoing integration, and which are therefore *rearoused and made plastic*, so as to be *reconsolidated in a new integration of memory and feeling*, emotionally simplified—rendered less conflictual and traumatic—by the experience of the dream.

Thus we can construe Freud's analysis as providing an example of the 'active systems' account of memory consolidation and reconsolidation. In this we see three stages:

1. First stage in the consolidation of memory: Freud's memories from the day—the penetration of Otto's remark, Freud's annoyance, his self-justificatory ruminations, and work late into the night—are transferred in SWS from the hippocampus to the cortical engram complexes that gave the events of the day their emotional significance, and with which they are to be integrated as long-term memories.

2. Second stage in the consolidation of memory: This transfer *rearouses* the target complexes, so that they can be *reconsolidated* in integration with the memories of the recent events they influenced. These older and deeper memories—Freud's advocacy of cocaine, his role in the deaths of his patient and friend, etc.—involve guilt, shame, and internal threats of punishment for his role in respect of toxic injections. Under the impact of this dual arousal of memory and (aversive/punishing) emotion, SWS gives way to REM and dreaming.

3. The final stage in the consolidation of memory: Freud's generative model responds to this arousal of aversive memory and emotion by producing the series of imaginary experiences of the dream, culminating in the wish fulfilling experience in which Freud reproaches Otto with **Injections of that sort ought not to be made so thoughtlessly.**

This powerfully rewarding experience of innocence and vindication, intensified by the embodiment of deep symbolic sexual significance,[24] leaves the cortically stored 'what, when, and where' aspects of Freud's memory intact. But it simplifies the complex of which these memories are part by altering their emotional registration in the amygdala. The fictive experiences of the dream thus freed Freud from the depressive and self-justificatory spiral into which he had fallen the day before, providing a kind of *overnight therapy* (Walker and van der Helm 2009), now being elaborated in neuroscientific work on dreaming and depression.[25] As suggested both by traditional psychoanalytic work on symbol formation (Spillius et al. 2011: Pt. 1, sec. 10) and recent work in cognitive neuroscience (Lewis et al. 2018), this can be seen as part of a more pervasive impact of the creative dreaming imagination on waking life and thought.

Freud's dream thus enabled him to address the difficulties he had felt to be accumulating in his practice in a way that was at once more realistic and more creative than his writing out of Irma's case history the evening before. On waking he turned from his sterile preoccupation with trying to justify his past practice to the fruitful one of seeking to understand his role in that practice at a deeper level. This began with his creative activity in analysing his dream, and so engaging consciously with the memories, emotions, and conflicts (particularly with his own superego) that had been roused by Otto's remark. In this he began the deep revisions in his theories and practice that showed weeks later in the *Project*, and that would eventually enable him to address the difficulties of his patients in a more thoughtful way than he had managed with Irma herself—and in particular by taking into fuller account the inner world of their psychic reality, as well as the outer world of their frustration and trauma.

Conclusion

Multiple pathways in contemporary scientific investigation of mind and brain lead to the relations of dreaming, memory, and emotion, and to the transformations that memories undergo in the processing of dreams. This consilience with psychoanalytic findings

indicates the value of collaboration, particularly among psychoanalysis, attachment, and neuroscience. The discoveries about engram complexes discussed in this chapter are already being applied to psychotherapy (Elsey and Kindt 2017). Their application to human dreaming requires further discovery and examination of the emotionally significant memories that are reconsolidated in our dreams.

Freud's method of analysing dreams by systematically comparing their manifest contents with the results of free association makes such discovery and examination possible, and there is no evidence that any other method can yet do so. So progress in this area, for neuroscience as well as the fields of psychology with which it is undergoing integration, may depend on the willingness of non-psychoanalytic researchers to explore the path originally broken by Freud.[26] If others do persevere in this—which is by no means certain, however great the potential benefit—they may come to see that path as leading to a greater understanding of the unconscious working of the FE-minimizing mind and brain.

References

Alberini, C. and Travaglia, A. (2017). 'Infantile Amnesia: A Critical Period of Learning to Learn and to Remember'. *Journal of Neuroscience37*(24): 578–579.

Alcaro, A., Carta, S., and Panksepp, J. (2017). 'The Affective Core of the Self: A Neuro-Archetypical Perspective on the Foundations of Human (and Animal) Subjectivity. *Frontiers in Psychology 8*: 1424. doi:10.3389/fpsyg.2017.01424

Alcaro, A. and Panksepp, J. (2011). 'The SEEKING mind: Primal neuro-affective substrates for appetitive incentive states and their pathological dynamics in addictions and depression'. *Neuroscience and Biobehavioral Reviews 35*: 1805–1820. doi:10.1016/j.neubiorev.2011.03.002

Ali, J., Spence, C., and Bremner, A. (2015). 'Human infants' ability to perceive touch in external space develops postnatally'. *Current Biology 25*: R965–R979.

Anda, R., Felitti, V., Bremner, J., Walker, J., Whitfield C., Perry B., Dube, S., and Giles, W. (2006). 'The enduring effects of abuse and related adverse experiences in childhood. A convergence of evidence from neurobiology and epidemiology'. *European Archives of Psychiatry and Clinical Neuroscience 256*: 174–186.

Atzila, S., Touroutogloua, A., Rudya, T., Salcedoa, S., Feldman, R., Hookera, J., Dickerson, B., Catana C., and Barrett, L. (2017). 'Dopamine in the medial amygdala network mediates human bonding'. *PNAS 114*(9): 2361–2366.

Baillargeon, R., Spelke, E., and Wasserman, S. (1985). 'Object permanence in five-month-old infants'. *Cognition 20*: 191–208.

Baillargeon, R., Stavans, M., Wu, D., Gertner,Y., Setoh, P., Kittredge, A., and Bernard, K. (2012). 'Object Individuation and Physical Reasoning in Infancy: An Integrative Account'. *Language, Learning and Development 8*(1): 4–46. doi:10.1080/15475441.2012.630610

Barrett, L. F. (2017). *How Emotions are Made*. London: MacMillian.

Beebe, B. and Lachman, F. (2014). *The origins of attachment: Infant research and adult treatment*. New York/Hove: Routledge.

Bion, W. (1962). 'A theory of thinking'. *International Journal of Psycho-Analysis 43*: 306–310 (reprinted in *Second Thoughts* (1967), London: Heinemann, pp. 110–119).

Bion, W. R. (1963). *Elements of Psycho-Analysis*. London: Heinemann.

Bivens, L. and Panksepp, J. (2011). *The Archaeology of the Mind*. New York: W. W. Norton.

Boccio, M., Nabavi, S., and Capogna, M. (2017). 'Synaptic Plasticity, Engrams, and Network Oscillations in Amygdala Circuits for Storage and Retrieval of Emotional Memories'. *Neuron 94*: May 17.

Bower, T. (1977). *Development in Infancy*. San Francisco, CA: W. H. Freeman.

Blatt, S. (2004). *Experiences of Depression: Theoretical, Clinical, and Research Perspectives*. Washington, DC: American Psychological Association.

Buzsaki, G. (2011). *Hippocampus. Scholarpedia 6*(1): 1468.

Campos, J., Barret, K., Lamb, M., Goldsmith, H., and Stenberg, C. (1983). 'Socioemotional Development'. In P. Mussen (ed.), *Handbook of Child Psychology*, Vol 3. New York: John Wiley.

Carhart-Harris, R. and Friston, K. (2010). 'The default-mode, ego-functions and free-energy: A neurobiological account of Freudian ideas'. *Brain 133*: 1265–1283.

Cartwright, R. (2010). *The Twenty-Four Hour Mind: The Role of Sleep and Dreaming in Our Mental Lives*. Oxford: Oxford University Press.

Clark, A. (2016). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*, Oxford/New York: Oxford University Press.

Conant, R. and Ashby, W. R. (1970). 'Every good regulator of a system must be a model of that system'. *International Journal of Systems Science 1*(2): 89–97.

Craig, A. D. (2015). *How Do You Feel: A Moment with your Neurobiological Self*. Princeton, NJ: Princeton University Press.

Craik. K. (1943). *The Nature of Explanation*. Cambridge: Cambridge University Press.

Crespi, B. and Semeniuk, C. (2004). 'Parent-Offspring Conflict in the Evolution of Vertebrate Reproductive Mode'. *The American Naturalist 163*(4).

Dahan, L., Astier, B., Vautrelle, N., Urbain, N., Kocis, B., and Chouvet, G. (2007). 'Prominent burst firing of dopaminergic neurons in the ventral tegmental area during paradoxical sleep'. *Neuropsychopharmacology 32*: 1232–1241. doi:10.1038/sj.npp. 1301251

Damasio, A. and Carvalho, G. (2013). 'The nature of feelings: evolutionary and neurobiological origins'. *Nature Reviews Neuroscience 14*: 143–152. doi:10.1038/nrn3403

Darwin, C. (1871). *The Descent of Man, and Selection in Relation to Sex*. London: John Murray.

Dayan, P., Hinton, G., Neal, R., and Zemel, R.(1995). 'The Helmholtz Machine'. *Neural Computation 7*: 899–904.

Dayan, P. and Abbot, L. (2001). *Theoretical Neuroscience*. London: MIT Press.

Douven, Igor (2011). 'Abduction'. In **The Stanford Encyclopedia of Philosophy** (Spring 2011 edn), E. N. Zalta (ed.).

Dudai, Y., Karni, A., and Born, J. (2015). 'The Consolidation and Transformation of Memory'. *Neuron 88*: October 7.

Elsey, J. and Kindt, M. (2017). 'Tackling maladaptive memories through reconsolidation: From neural to clinical science'. *Neurobiology of Learning and Memory 142*: 108–117.

Enlow, M., Egeland, B., Carlson, E., Blood, E., and Wright, R. (2014). 'Mother–infant attachmentand the intergenerational transmission of post traumatic stress disorder'. *Developmental Psychopathology 26*: 41–65. doi:10.1017/S0954579413000515

Feld, G. and Born, J. (2017). 'Sculpting Memory in Sleep'. *Current Opinion in Neurobiology 44*: 20–27.

Fischmann, T., Russ, M. and Leuzinger-Bohleber, M. (2013). 'Trauma, dream, and psychic change in psychoanalyses: A dialog between psychoanalysis and the neurosciences'. *Frontiers in Human Neuroscience 7*: 877. doi:10.3389/fnhum.2013.00877

Fotopoulou, A. and Tsakiris, M. (2017). 'Mentalizing homeostasis: The social origins of interoceptive inference'. *Neuropsychoanalysis 19*: 3–29. doi: 10.1080/15294145.2017.1294031

Freud, S. (1895/1950). 'Project for a Scientific Psychology'. In J. Strachey (ed.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 1 (pp. 295–324). London: Hogarth Press.

Freud, S. (1900/1953). *The Interpretation of Dreams*. In J. Strachey (ed.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vols 4 and 5. London: Hogarth Press.

Freud, S. (1911b/1957). 'Formulations on the Two Principles of Mental Functioning'. In J. Strachey (ed.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 12 (pp. 1–628). London: Hogarth Press.

Freud, S. (1915–17/1957). 'Introductory Lectures on Psycho-Analysis'. In J. Strachey (ed.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 15 (pp. 218–227). London: Hogarth Press.

Freud, S. (1920/1955). 'Beyond the Pleasure Principle'. In J. Strachey (ed.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 18 (pp. 1–64). London: Hogarth Press.

Freud, S. (1921/1955). 'Group Psychology and the Analysis of the Ego'. In J. Strachey (ed.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 18 (pp. 67–143). London: Hogarth Press.

Freud, S. (1923/1955). 'The Ego and the Id'. In J. Strachey (ed.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 19 (pp. 1–60). London: Hogarth Press.

Friston, K. (2010). 'The free-energy principle:a unified brain theory?'. *Nature Reviews Neuroscience 11*: 127–138. doi:10.1038/nrn2787

Friston, K. (2012). 'Prediction, perception and agency'. *Internatio nal Journal of Psychophysiology 83*: 248–252. doi:10.1016/j.ijpsycho.2011.11.014

Friston, K. (2013). 'Life as we know it'. *Journal of the Royal Society Interface 10*: 0475. doi:10.1098/rsif.2013.0475

Genzel, L., Spoormaker, V., Konrad, M., and Dresler, M. (2015).'The role of rapid eyemovement sleep for amygdala-related memory processing'. *Neurobiology of Learning and Memory 122*: 110–121. doi:10.1016/j.nlm.2015.01.008

Helmholtz, H. (1865/1924). *A Treatise on Physiological Optics*, Vol. III, trans. J. P. C. Southall. New York: Optical Society of America.

Hobson, A., Hong C., and Friston, K. (2014). 'Virtual reality and consciousness inference in dreaming'. *Frontiers in Psychology 5*: 1133. doi:10.3389/fpsyg.2014.01133

Hobson, A. (2014). *Ego Damage and Repair:Towards a Psychodynamic Neurology*. London: Karnac.

Hobson, A. (2015). *Psychodynamic Neurology: Dreams,Consciousness, and Virtual Reality*. New York: Taylor and Francis.

Hohwy, J. (2013). *The Predictive Mind*. Oxford: Oxford University Press.

Hopkins, J. (1987). 'Synthesis in the Imagination: Psychoanalysis, Infantile Experience, and the Concept of an Object'. In J. Russell (ed.), *Philosophical Perspectives on Developmental Psychology*. Oxford: Basil Blackwell.

Hopkins, J. (1999). 'Patterns of interpretation: Speech, action, and dream'. In L. Marcus (ed.), *Cultural Documents: The Interpretation of Dreams* (pp. 123–169). Manchester: Manchester University Press.

Hopkins, J. (2002). 'Psychoanalysis, Metaphor, and the Concept of Mind'. In M. Levine (ed.), *The Analytic Freud* (pp. 22–47). London: Routledge.

Hopkins, J. (2012). 'Psychoanalysis Representation and Neuroscience: The Freudian unconscious and the Bayesian brain'. In Fotopolu, Pfaff, and Conway (eds), *From the Couch to the Lab: Psychoanalysis, Neuroscience and Cognitive Psychology in Dialogue*. Oxford: Oxford University Press.

Hopkins, J. (2018/2014). 'Kantian Neuroscience and Radical Interpretation: Ways of Meaning in the Bayesian Brain'. In G/ Oritz-Milan and J. Parero (eds), *Mind, Language, and Morality*, London: Routledge.

Hopkins, J. (2015). 'The Significance of Consilience: Psychoanalysis, Attachment, Neuroscience, and Evolution'. In S. Boag, L. Brakel, and V. Talvete (eds), *Psychoanalysis and Philosophy of Mind: Unconscious mentality in the 21st century*. London: Karnac.

Hopkins, J. (2016). 'Free Energy and Virtual Reality in Psychoanalysis and Neuroscience: A Complexity Theory of Mental Disorder'. *Frontiers in Psychology: Frontiers in Cognitive Science 7*: 922.

Hopkins, J. (2018a). 'Darwin, Freud, and Group Conflict'. In K. Figlio and S. Krueger (ed.), *Fomenting Political Violence* London: Polity Press.

Hopkins, J. (2018b). 'Freud and Mental Health'. In P. Adamson (ed.), *The Oxford Handbook of the Philosophy of Health*. Oxford: Oxford University Press.

Horton, C. L. and Malinowski, J. (2015). 'Autobiographical memory and hyperassociativity. *Frontiers in Psychopathology 6*: 874. doi:10.3389/fpsyg.2015.00874

Howe, G. (2011). *Attachment Across the Lifecourse*. New York: Palgrave Macmillan

Johnson, S., Amso, D., and Slemmer., J. (2003). 'Development of object concepts in infancy: Evidence for early learning in an eye-tracking paradigm'. *PNAS 18*(100): 10568–10573.

Johnson-Laird, P. (1983). *Mental Models*. Cambridge: Cambridge University Press.

Kant, I. (1764/2007). 'Essay on the maladies of the head'. In P. Guyer and H. Wood, *The Cambridge Edition of the Works of Immanuel Kant in Translation: Anthropology, History, and Education*, trans. H. Wilson. Cambridge: Cambridge University Press.

Kaufman, J., Csibra, G., and Johnson, M. H. (2005). 'Oscillatory activity in the infant brain reflects object maintenance'. *Proceedings of the National Academy of Science USA 102*: 15271–15274. doi:10.1073/pnas.0507 626102

Klein, M. (1952/1972). 'The Origins of Transference'. In *Envy and Gratitude and Other Works 1946–1953*. London: Hogarth Press.

Kumin, I. (1996). *Pre-object Relatedness: Early Attachment and the Psychoanalytic Situation*. New York and London: Guilford Press.

Kurth, S., Olini, N., Huber, R. et al. (2015). 'Sleep and Early Cortical Development'. *Current Sleep Medicine Reports 1*(1): 64–73. doi:10.1007/s40675-014-0002-8

Latchoumane, C., Ngo, H., Born, J., and Shin, H. (2017). 'Thalamic Spindles Promote Memory Formation during Sleep through Triple Phase-Locking of Cortical, Thalamic, and Hippocampal Rhythms'. *Neuron 95*(2): 442–435. doi:10.1016/j.neuron.2017.06.025

LeDoux, J. (2015). *Anxious: The Modern Mind in the Age of Anxiety*. London: Oneworld Publishers.

LeDoux, J. and Brown, R. (2017). 'A higher-order theory of emotional consciousness'. *PNAS 114*(10). doi:10.1073/pnas.1619316114

Leung, S., Mareschal, D., Rowsell, R., Simpson, D., Iaria, L., and Kaufman, J. (2017). 'Oscillatory Activity in the Infant Brain and the Representation of Small Numbers'. *Frontiers in Systems Neuroscience: 10*(4).

Levin, R. and Nielsen, T. (2009). 'Nightmares, bad dreams, and emotion dysregulation'. *Current Directions in Psychological Science 18*: 84–88. doi:10.1111/j.1467–8721.2009.01614.x

Lewis, P., Knoblich, G., and Poe, G. (2018). 'How memory replay in sleep boosts creative problem-solving'. *Trends in Cognitive Sciences 22*: 6. doi:10.1016/j.tics.2018.03.009

Li, W., Ma, L., Yang, G., and Gan, W.-B. (2017). 'REM sleep selectively prunes and maintains new synapses in development and learning'. *Nature Neuroscience 20*: 427–437. doi:10.1038/nn.4479

Malinowski, J. and Horton, C. (2014a). 'Emotion but not stress modulates the incorporation of waking experiences into dreams'. *Dreaming 24*: 18–31. doi:10.1037/a0036017

Malinowski, J. and Horton, C. (2014b). 'Memory sources of dreams: The incorporation of autobiographical rather than episodic experiences'. *Journal of Sleep Research 23*: 441–447. doi:10.1111/jsr.12134

Malinowski, J. and Horton, C. (2015). 'Metaphor and hyperassociativity: The imagination mechanisms behind emotion assimilation in sleep and dreaming'. *Frontiers of Psychology 6*: 1132. doi:10.3389/fpsyg.2015.01132

McCrory, E., Gerin, M., and Viding, E. (2017). 'Annual Research Review: Childhood maltreatment, latent vulnerability and the shift to preventative psychiatry—the contribution of functional brain imaging'. *Journal of Child Psychology and Psychiatry 58*(44): 338–357.

Metzinger, T. and Weise, W. (2017). *Philosophy and Predictive Processing*. Frankfurt am Main: MIND Group.

Mosquera, D., Gonzales, A., and Leeds, A. (2014). 'Early experience, structural dissociation, and emotional dysregulation in borderline personality disorder: The role of insecure and disorganized attachment'. *Bordeline Personality Disorder. Emotional Dysregulation 1*: 15. doi:10.1186/2051-6673-1-15

Panksepp, J. (1998). *Affective Neuroscience*. Oxford: Oxford University Press.

Parsons, C., Young, K., Joensson, M., Brattico, E., Hyam, H., Stein, A., Green, A., Aziz, T., and Kringlebach, M. (2014). 'Ready for action: A role for the human midbrain in responding to infant vocalizations'. *SCAN 9*: 977–984.

Piaget, J. and Inhelder, B. (1969). *The Psychology of the Child*. London: Routledge and Kegan Paul.

Pellouzio, G., Rigoli, F., and Friston, K. (2015).'Active Inference, homeostatis regulation, and adaptive behavioural control'. *Progress in Neurobiology 134*: 17–35. doi:10.1016/j.pneurobio.2015.09.001

Ramirez, S., Liu, X., MacDonald, C., Moffa, A., Zhou, J., Redondo, R., and Tonegawa, S. (2015). 'Activating positive memory engrams suppresses depression-like behaviour'. *Nature 522*: 336–339.

Ramstead, M., Badcock, P, and Friston, K. (2017). 'Answering Schroedinger's Question: A Free-Energy Formulation'. *Physics of Life Review 24*: 1–16. doi:10.1016/j.plrev.2017.09.001

Rasch, B. and Born, J. (2013). 'About sleep's role in memory'. *Physiological Reviews: American Journal of Physiology 93*: 681–766.

Reissland, N., Francis , B., Mason, J. and Lincoln, K (2011). 'Do Facial Expressions Develop before Birth?'. *PLoS ONE 6* (8): e24081. doi:10.1371/journal.pone.0024081

Reissland, N., Francis, B., and Mason, J. (2013). 'Can Healthy Fetuses Show Facial Expressions of "Pain" or "Distress"?'. *PLoS ONE 8*(6): e65530. doi:10.1371/ journal.pone. 0065530

Saarimäki, H., Gotsopoulos, A., Jääskeläinen, I., Lampinen, J., Vuilleumier, P., Hari, R., Sams, M., and Nummenmaa, L. (2016). 'Discrete Neural Signatures of Basic Emotions'. *Cerebral Cortex 26*(6): 2563–2573. doi:10.1093/cercor/bhv086

Schaffer, H. R. (1971). *The Growth of Sociability*. Harmondsworth: Penguin Books.

Segal, H. (1964). 'Symposium on Phantasy—Phantasy and Other Mental Processes'. *International Journal of Psycho-Analysis 45*: 191–194.

Segal, H. (1975/1981). 'A psycho-analytic approach to the treatment of schizophrenia'. In *The Work of Hanna Segal* (pp. 131–136). New York: Jason Aronson.

Seth, A. (2015). 'The Cybernetic Bayesian Brain'. In T. Metzinger and J. M. Windt (eds), *Open MIND*: 35(T). Frankfurt am Main: MIND Group. doi:10.15502/9783958570108

Seth, A. and Friston, K. (2016). 'Active interoceptive inference and the emotional brain'. *Philosophical Transactions of the Royal Society 371* (1708). doi:10.1098/rstb.2016.0007

Solms, M. (2013). 'The conscious Id'. *Neuropsychoanalysis 15*(1).

Solms, M. and Panksepp, J. (2012). 'The "Id" Knows More than the "Ego" Admits: Neuropsychoanalytic and Primal Consciousness Perspectives on the Interface Between Affective and Cognitive Neuroscience'. *Brain Science 2*: 147–175. doi:10.3390/ brainsci2020147

Solms, M. and Turnbull, O. (2002). *The Brain and the Inner World: An Introduction to the Neuroscience of Subjective Experience*. London: Karnac.

Solomon, J. and George, C. (1999). *Attachment Disorganization*. New York: Guilford Press.

Spillius, E., Milton, J., Garvey, P., Couve, C., and Steiner, D. (2011). *The New Dictionary of Kleinian Thought*. Hove: Routlege.

Stenberg, C. R., Campos, J., and Emde, R. N. (1983). 'The facial expression of anger in seven-months-old infants'. *Child Development 54*: 178–184.

Stenner, K. (2010). *The Authoritarian Dynamic*. Cambridge: Cambridge University Press

Swanson, L. R. (2016). 'The predictive processing paradigm has roots in Kant'. *Frontiers in Systems Neuroscience 10*: 79. doi:10.3389/fnsys.2016.00079

Tappenberg, T. (2010). *Foundations of Computational Neuroscience*. Oxford: Oxford University Press.

Tonegawa, S., Liu, X., Ramirez, S., and Redondo, R. (2015a). 'Memory Engram Cells Have Come of Age'. *Neuron 85*: 918–931.

Tonegawa, S., Pignatelli, M., Roy, D., and Ryan, T. (2015b). 'Memory Engram Storage and Retrieval'. *Current Opinion in Neurobiology 35*: 101–109.

Tononi, G. and Cirelli, C. (2014). 'Sleep and the price of plasticity: From synaptic and cellular homeostasis to memory consolidation and integration'. *Neuron 81*: 12–34. doi: 10.1016/j.neuron.2013.12.025

Travaglia, A., Bisazi, R., Sweet, E., Blitzer, R., and Alberini, C. (2016). 'Infantile amnesia reflects a developmental critical period for hippocampal learning'. *Nature Neuroscience 19*: 1225–1233.

Vogel, G., McAbee, R. J., Barker., K., and Thurmond, R. (1977). 'Endogenous Depression Improvement and REM Pressure'. *Archives of General Psychiatry 34*: 96–97.

Volkow, N., Wise, R., and Baier, R. (2017). 'The dopamine motive system: Implications for drug and food addiction'. *Nature Reviews Neuroscience 18*(12): 741–752. doi:10.1038/nrn.2017.130

Walker, M. and van der Helm, E. (2009). 'Overnight Therapy? The Role of Sleep in Emotional Brain Processing'. *Psychological Bulletin 135*(5): 731–748.

Watt, D. and Panksepp, J. (2009). 'Depression: An evolutionarily conserved mechanism?'. *Neuropsychoanalysis 2*(1): 93.

Wilcox, T. (1999). 'Object individuation: Infants' use of shape, size, pattern, and color'. *Cognition 72*: 125–166.

Wilcox, T. and Biondi, M. (2015). 'Object processing in the infant: Lessons from neuroscience'. *Trends in Cognitive Sciences 19* (7): 406–413.

Notes:

(1.) Thus as briefly discussed later below the psychoanalytic mechanisms of identification and projection seem clearly to play a role in the ingroup cooperation for outgroup conflict that Darwin and contemporary theorists of culture gene co-evolution regard as formative for our species, and both the Oedipus complex and the death drive also seem clearly related to parent–offspring, sexual, and sibling conflict. These are discussed together in Hopkins (2018a); and Hopkins (2015) discusses the relations of psychoanalysis, evolution, attachment, and neuroscience more generally. The integration of Freudian and contemporary free energy neuroscience in the complexity theory of mental disorder sketched later on is treated at greater length in Hopkins (2016).

(2.) The programme has a (technical) website at https://en.wikipedia.org/wiki/Free_energy_principle.

(3.) On generative models in contemporary neuroscience, see Ch. 10 of Dayan and Abbot (2001); for an exegesis of Friston's use, see Ch. 10.3ff of Tappenberg (2010).

(4.) This account was ignored by the founders of twentieth-century analytical philosophy. Succumbing to what Wittgenstein later described as 'the picture that forces itself on us at every turn', Frege, Russell, and the early Wittgenstein turned their backs on Helmholtz's conception of perceptual experience as synthesized by the brain. Instead they adopted a Cartesian conception in which such experiences were the private 'sense-data' upon which all knowledge, including the existence of bodies and brains, had somehow to be based. See Hopkins (2014/2018c) for an account of the development of Helmholtz's ideas concerning modelling through Frege and the early Wittgenstein, which ideas prompted Craik's (1943) account of mental models, later taken up in cognitive science, e.g. in Johnson-Laird (1983).

(5.) This holistic interpretive integration of the data of free association is discussed in some detail in Hopkins (1999). Although other forms of investigation of the emotional functions of dreaming increasingly acknowledge the evidential importance of the dreamer's associations and memories (see, e.g. Malinowski and Horton 2015), no mode of study apart from psychoanalysis has taken them so thoroughly into account. If as argued later dreaming functions as part of the emotional synthesis of long-term memory then progress in understanding both dreaming and memory will partly depend upon other disciplines following this lead.

(6.) In arriving at this account Freud started to provide detailed empirical support for a tradition of thought about dreaming and disorder that went back to Plato, and was summarized in Kant's (1764/2007: 71) claim that 'the deranged person' was 'a dreamer in waking'.

(7.) While still a medical student Freud was invited by the celebrated physiologist Ernest Bruke to conduct neurological research in his laboratory. Prior to practising as a psychiatrist he published over a hundred papers, as well as monographs on disorders of movement and childhood cerebral palsy that established him as an expert in these fields.

(8.) Helmholtz's notion of thermodynamic free energy measures a physical quantity, whereas the notion of variational free energy (FE) provides a measure of Bayesian model evidence that is computable from parameters of a model. Similarities and differences between the notions are discussed in some detail at https://en.wikipedia.org/wiki/Free_energy_principle. The two are related mathematically, since the equations describing variational FE derive from Helmholtz's describing thermodynamic free energy; and for this reason they share the terminology of statistical physics. Also the complexity term of variational free energy shares the same fixed point as Helmholtz free energy for a system that is thermodynamically closed but not isolated.

(9.) 'Optimally': that is, by the criteria imposed by evolution, which would determine the prior hypotheses and major modes of processing of the model, as well as by the environment, including experiential learning over the lifespan. Hence only in some areas (e.g. perception) does the brain appear as operating with Bayesian rationality. In others (e.g. the assignment of properties to members of idealized ingroups as opposed to derogated outgroups) it must be understood as operating in accord with priors imposed by evolution (in this case those that sustain the process of ingroup cooperation for outgroup conflict, 'the competition of tribe with tribe', that Darwin thought 'sufficed to raise man to his present high position in the organic scale' (1871: 157)). In enabling us to understand such oppositions between the imperatives of rationality and evolution the FE programme encourages enquiry that parallels Quine's conception of the naturalization of epistemology.

(10.) Since the turn of the century Friston and others have directed this approach to the rest of organic life (Friston 2013; Ramstead et al. 2017) and applied it to the brain in rapidly increasing scope and detail. This has established the minimization of FE as a paradigm for neuroscience, described variously as predictive coding, predictive processing, the predictive brain, and the predictive mind. And since the probabilistic regulatory generative models envisaged in this approach would encompass a host of others—the mental models of cognitive science, the internal working models of attachment theory, the internal objects of psychoanalysis, etc.—the programme naturally extends through psychology as well.

Although Friston's own presentations are often forbiddingly technical, the main ideas have been given clear philosophical expositions in Hohwy (2013) and Clark (2016), as well as a recent collection of essays edited by Metzinger and Weise (2017). Seth (2015) is a good introductory starting point, and emphasizes the role of interoception in the generation of emotion, which is discussed later in this chapter.

(11.) This continuity contradicts claims that human thought cannot be understood as having arisen via natural selection. And although specifics of Friston's account of the brain's model are controversial, the existence of some such model (map, etc.) is widely accepted in neuroscience. The brain viewers at http://gallantlab.org/index.php/brain-viewer/ display a striking range of information about the modelling of the external world in the cortex (see also the video introduction at https://www.youtube.com/watch?v=u9nMfaWqkVE).

(12.) Friston's collaborator Alan Hobson, for decades a zealous neuroscientific critic of Freud, has now published two books (Hobson 2014, 2015) applying the virtual reality/generative model conception to link dreaming with mental disorder, although via neurotransmission rather than Freudian phantasy. While continuing to criticize Freud, Hobson now stresses that his own work 'takes up the *Project for a Scientific Psychology* exactly where Freud left it in 1895' (Hobson 2015: 5).

(13.) These systems are described differently by different investigators. For example, LeDoux (2015) describes them as 'survival circuits', and Volkow et al. (2017) describe the SEEKING system as the Dopamine Motive System. The most detailed recent delineation is Panksepp's (1998), which has been updated for therapists and general readers in Bivens and Panksepp (2011).

(14.) One important difference is that Freud did not anticipate the current claim that the brainstem systems generate consciousness as well as emotion, both of which are elaborated in the cortex. For this see Solms and Panksepp (2012). Again many authors, such as Barrett (2017) and LeDouz and Brown (2017) argue that the contextual cortical elaboration of emotion is inconsistent with the notion of basic emotions. But as Panksepp argues, subcortical generation of the basics is consistent with any degree of cortical elaboration of the final states, including that emphasized by Barrett; and as noted by Saarimäki et al. (2016) the subcortical generators have 'fingerprints' in their complex cortical derivatives.

(15.) This is the kind of dialectic that Bion (1962: 30) attempts to describe phenomenologically as follows:

> The infant suffers pangs of hunger and feels it's dying; racked by guilt and anxiety and impelled by greed, it messes itself and cries. The mother picks it up, feeds it and comforts it and eventually the infant sleeps. In forming the model to represent the feelings of the infant, we have the following version: the infant, filled with painful lumps of faeces, guilt, fears of impending death, chunks of greed, meanness and urine, evacuated these bad objects into the breast that is not there. As it does so, the good object turns the no-breast (mouth) into a breast, the faeces and urine into milk, the fears of impending death and anxiety into validity and confidence, the greed and meanness into feelings of love and generosity; and the infant sucks in its bad property, now translated into goodness, back again.

(16.) Some of main developments occur in and between the 'what' and 'where' object-processing pathways (Wilcox and Bondi 2015), with an early representation of an identified object as lasting while unobserved *perhaps* signalled at six months by gamma wave activation in the right temporal cortex (Kaufman et al. 2005; see also Leung et al. 2017).

This is consistent with an experimental tradition that at first focused on the mother as object of perception, but later on other (emotionally less significant) objects. Thus Bower (1977) reported an experiment with mirrors (discussed in Hopkins 1987) that cohered with Klein's dating of the depressive position.

> If one presents the infant with multiple images of its mother—say three 'mothers'—the infant of less than five months is not disturbed at all but will in fact interact with all three 'mothers' in turn. If the setup provides one mother and two strangers, the infant will preferentially interact with its mother and still show no

signs of disturbance. However, past the age of five months (after the coordination of place and movement) the sight of three 'mothers' becomes very disturbing to the infant. At this same age a setup of one mother and two strangers has no effect. I would contend that this in fact shows that the young infant (less than five months old) thinks it has a multiplicity of mothers, whereas the older infant knows it has only one

(Bower 1977: 217).

Although this experiment has not been repeated in contemporary conditions, it was consistent with experiments on occlusion (e.g. Baillargeon et al. 1985; Wilcox 1999) and eye-tracking (Johnson et al. 2003) done afterwards. Later work (Baillargeon et al. 2012) has clarified how the infant initially uses parameters of shape and then others as the ability to individuate and track objects is refined.

(17.) See also Klein (1952/1972: 54). For comparison of Piaget and Klein on infantile experience and the object concept see Hopkins 1987.

(18.) Thus as reported in (Campos et al. 1983), a three-month-old infant made angry by someone using her hand to impede movement will express anger *at the impeding hand*. By contrast a seven-month-old infant made angry in the same way expresses anger *not at the hand, but at the face*. Apparently at three months the infant has not, whereas by seven months it has, integrated episodic experiences with the mother into an image of an anatomically whole and enduring person. Also by seven months, as reported in (Stenberg et al. 1983) this infant's anger is regulated by experience of the persons she knows. If the infant is annoyed twice by mother, or again twice by a stranger, the infant is angry both times. But an infant annoyed *first* by a stranger and *then* by mother is especially angry, indicating that the infant's expectation that mother will comfort her after intrusions by a stranger has been betrayed.

(19.) The work reported in Ali et al. (2015) suggests that infants' sense of bodily touch becomes anchored to objects in the surrounding space at four to six months. And recently Seth and Friston (2017) have suggested that failure in integrating interoceptive and exteroceptive sensory information—in particular, failure properly to integrate 'interoceptive signals associated with nurturance (e.g. breastfeeding) during affiliative interactions with (m)others' may render autistic infants incapable of learning 'that the nursing and prosocial mother were the same hidden cause or external object' (Seth and Friston 2017: 7, and Fig. 4). If this is correct it would place autism among the disorders that result from failure to negotiate this depressive-position integration.

(20.) For a fuller discussion of these developments, which also include the regulation of parent–offspring conflict and sibling rivalry as part of the Oedipus complex, see Hopkins (2018b).

(21.) For further discussion and citations see the section on 'Waking Disorder and Complexity Reduction in Sleep' in Hopkins (2016) as well as the work on dreaming and depression cited later in this section.

(22.) For a fuller discussion of this dream which bears out claims made in passing here see Hopkins (2015), and for further methodological discussion see Hopkins (1999).

(23.) See also Walker and van der Helm (2009) and Cartwright (2010).

(24.) On symbolism see Petocz, this volume. For the symbolism in this dream see the discussions in Hopkins (2015), and Hopkins (2002) on psychoanalysis and conceptual metaphor, which is also discussed in the work by Malinowski and Horton cited here.

(25.) See the discussion in Cartwright (2010: ch. 4), following up the work Vogel et al. (1977); and the continuation of Cartwright's work in the neuropsychoanalytic research project on dreaming and depression described in Fischmann et al. (2013).

(26.) Thus for example the main arguments and claims of a recent series of non-psychoanalytic papers by Malinowski and Horton (e.g. 2014a, 2014b, 2015; see also Horton and Malinowski 2015) are consilient both with Freud and with the argument of the present paper. They discuss role of memory, the dreamer's associations, and symbolism, and see dreams as performing 'emotion assimilation' via symbolic embodied cognition that ameliorates the 'emotional intensity' of memories transformed in them.

Jim Hopkins

Department of Philosophy, King's College London

Oxford Handbooks Online

Abstract and Keywords

This chapter reviews the debate between 'realist' and 'constructivist' understandings of the psychoanalytic unconscious. To oversimplify, realists hold that unconscious mental states exist in the analysand's mind fully formed and with determinate intentional content, independent of consciousness, and these are discovered in analysis. Constructivists (including relationalists and intersubjectivists) hold that the unconscious meaning of clinical material does not exist 'preorganized' in the analysand's mind, but is constructed, not discovered, through the analytic relationship. The chapter argues that the debate is multiply confused. For example, different meanings of 'psychoanalysis' and 'constructivism' are at play, and a number of central arguments rest on misunderstandings of complex philosophical positions concerning the status of science and the nature of human knowledge. Once these confusions are removed, an understanding of the psychoanalytic unconscious that retains the strengths of both realism and constructivism presents itself.

Keywords: unconscious, realism, constructivism, relationalism, science, hermeneutics, Stephen Mitchell

Michael Lacewing

Introduction: Psychoanalysis and Theory

PSYCHOANALYTIC theories of the unconscious have changed greatly since Freud.[1] I shall review a debate that has taken place over the last forty years, predominantly in the USA, between two families of theories I shall term 'realist' and 'constructivist'. An oversimple first characterization of the two positions is this: realists hold that unconscious mental states exist fully formed and with determinate intentional content, independent of consciousness, and are part of the analysand's mind, playing a determinate role in their thoughts, feelings, and behaviour, and this is what is discovered in analysis. By contrast,

constructivists—by which I mean to include not only those who self-identify as 'constructivists', but also many relationalists and intersubjectivists—claim that the unconscious meaning of clinical material is constructed, not discovered, and this is of central importance to understanding psychoanalytic knowledge and the nature of the unconscious. They hold that not only are psychological structures *built* from patterns of interpersonal interaction (ontology), individuals can only be *understood* within an 'interactional field' (epistemology). It is only through interacting that people can come to understand one another (in analysis as elsewhere), and to this interaction, both bring (p. 408) their individual subjectivity. And so, says Stephen Mitchell (1998: 18), one of the originators of relationalism, we must reject the claims of classical Freudian practice 'that the central dynamics relevant to the analytic process are pre-organised in the patient's mind and that the analyst is in a privileged position to gain access to them'.

I shall argue that the debate is multiply confused, with detractors on both sides misunderstanding not only the claims and arguments of their opponents, but their own positions and commitments as well. Once we untangle the many threads of thought that have fed into the debate, this becomes less surprising. Part of my purpose is to explore the extent to which constructivism and realism may be compatible once the debate is made clearer.

One element of the debate—more often implicit than explicit—has concerned just what is meant by 'psychoanalysis'. As most commonly understood, psychoanalysis is a form of psychotherapy, the relationship between analyst and analysand in the clinical setting. We should grant that the heart of psychoanalysis is the clinical practice, including the observations, interpretations, and inferences about one individual's psychological experience within the unique relationship between this analysand and this analyst. But it is a mistake to restrict the meaning of 'psychoanalysis' to this practice. Psychoanalysis is equally a *theory* about the nature, development, and functioning of the human mind, especially in relation to motives (see also Eagle, this volume). There are a number of features of this theory, particularly concerning the role and nature of unconscious mental states and processes, that make it recognizably distinct and a competitor with other psychological theories deriving, for instance, from cognitive psychology or neuroscience. As we shall see, some of the confusion in the debate has derived from not clarifying whether claims being advanced concern the therapeutic practice, the psychological theory, or both.

To complicate matters further, the theoretical functioning of psychoanalysis takes place at many levels, or again in many contexts, and these are also not always clearly differentiated in the debate:

> **1.** At the clinical level, we may observe that psychoanalysts generalize over individual cases to make more general claims about the nature and function of particular psychological processes or structures manifest in clinical interaction, such as the meaningfulness of dreams, the occurrence of psychological defence, or the phenomenon of transference. We can call this 'low-level' or '*experience-near*'

theorizing. Such claims are so central to psychoanalysis they are very often not recognized as forming part of an explanatory psychological theory, yet they are taken to characterize human psychology universally (or at least widely within a culture), and so should be distinguished from claims made about individual experience and functioning.

2. Psychoanalysts frequently state or argue for claims about more abstract processes or structures, often without (apparently) inferring them from clinical data, although clinical data are frequently presented as supporting or illustrating such claims. Call this 'high-level' or '*abstract*' *theorizing*. One example would be Freud's (1923) (p. 409) theory of the superego or again general theories of the origin, structure, and influence of object relations, such as we find in Fairbairn (1952) or Kohut (1971).

3. They also range into more speculative discussion about human psychology, and in some cases, take up or begin from philosophical arguments and theories, e.g. Fromm (1979), Stern (1997), or Stolorow and Atwood (2016). Call this '*applied philosophizing*'. We will see that a good part of the debate over how to understand the unconscious has been informed by such theorizing.

4. On other occasions, for any of 1 to 3 above, psychoanalysts appeal to theories or empirical findings in other disciplines, from other branches of psychology to anthropology and linguistics to biology, from Freud's (1895) 'economic' theory to Mitchell's (1988) account of object relations. Call this '*interdisciplinary*' *theorizing* at the relevant level.

5. Finally, psychoanalysts apply psychoanalytic theory to wider issues, e.g. the nature of social and cultural practices, religion, group psychology, etc. Call this '*social*' *theorizing*.

In what follows, I shall occasionally refer back to these distinctions to diagnose a confusion or line of thought. But let us get on. What is the debate about, and how did it begin?

The Critique of Freud

In the debate, Freud is presented as holding a strongly realist position on the unconscious. Whether Freud was a realist in the sense described earlier is not my concern here. But we can note that, at different points in his career, Freud claims that hypnosis recovers memories; that the meaning(s) of a dream lie(s) in specific, determinate latent thoughts and wishes; that psychoanalytic interpretation identifies specific defences that have distorted intentional content in specific ways (e.g. through reversal, projection, etc.); that interpretation uncovers repressed wishes, which exist as such prior to lifting the repression; and so on. Importantly, these mental states are representations with determinate intentional content, to be understood in their nature and ontology along the lines of conscious mental states, though allowing for the differences imposed by primary process in the system *Ucs*. Of course, Freud also holds that there is great complexity in the unconscious, and so discovering exactly which unconscious mental states are expressed in a dream or operating in a defence can be very

challenging, but this is a result of myriad associations and primary process substitutions, rather than the result of any indeterminacy in the unconscious mental states themselves.

Freud is understood, then, as thinking of the unconscious as 'a storehouse of fully formed mental contents that, when uncovered, appear in consciousness, unchanged in their original form' (Eagle 2011: 112). But this model, Eagle continues, 'has been rejected in much of the contemporary psychoanalytic, cognitive psychology, and philosophical literature'. This rejection—at least within psychoanalysis—went hand in hand with a (p. 410) broader critique of Freud's metapsychology, and this broader critique was itself part of a reconsideration of whether psychoanalysis could properly be considered a science, and this question was in turn situated within yet broader philosophical concerns about the nature of human knowledge as such.

One result of this breadth and complexity in the debate has been that, in a great many of the most influential contributions over how to think of the unconscious, these issues have not been adequately separated. And a result of this, I believe, has been the imposition of an unhelpful dualistic, either-or framework—one *either* thinks:

> **1)** that psychoanalysis is and aspires to be a science (which is *the* paradigm of human epistemic achievement), *and*
>
> **2)** that metapsychology (of some kind) is central to this, *and*
>
> **3)** that realism about unconscious mental states is correct

or one thinks

> **1')** that psychoanalysis is not a science (which, in any case, is at best one form of knowing among others), *and*
>
> **2')** that metapsychology has little place, *and*
>
> **3')** that realism about the unconscious is false.

But why accept the alignment of these theoretical commitments? Very few writers have recognized even the possibility of holding, say, that psychoanalysis is some kind of science (which is one form of knowledge among others) but realism about the unconscious is false, or again that realism is true, but Freud's metapsychological and scientific aspirations were misplaced, or any of a number of other possible combinations. One purpose of this chapter is to free us from the straitjacket of thought that has been voluntarily but unwittingly donned by many contributors to the debate.

The rejection of Freud's model of the unconscious gained force and momentum in the 1960s, through critiques both internal and external to psychoanalysis. Such critique had always been around, but additional impetus—certainly in the USA—was sparked by a symposium arranged by Sydney Hook in 1958 in New York, at which Hook and Ernest Nagel, among others, presented forceful criticisms of psychoanalysis as science which Heinz Hartmann's reply did little to quash (see Hook 1959). Critics outside psychoanalysis engaged with issues of the truth of its claims in general and the reliability of its method, both in generating and in drawing inferences from the clinical data. At the

same time, the internal critique questioned both the validity and the usefulness of Freud's metapsychology and his drive theory in particular.

It was, in part, the theory of drives that supported Freud's claims of realism. Freud had argued that biologically based drives provide the most fundamental sources of psychological motivation. Each drive introduces energy into the psychological system that presses for action. It generates motivation that aims at the satisfaction of the drive, attaching to an object that can supply this, which helps determine the intentional (p. 411) content of the corresponding mental state. The development of the drives in childhood, in particular of the multifaceted libido, he argued, contributes materially to the development of the subject's mind and mental health or psychopathology.

This theory, together with a number of implications that Freud drew concerning clinical practice and the interpretation of clinical material, had drawn fire ever since its formulation and development around the turn of the twentieth century. It was hotly disputed not only by the scientific community, but also by a number of Freud's early collaborators and supporters, including Adler, Jung, Stekel, Forel, and Bleuler (see Borch-Jacobsen and Shamdasani 2012: ch. 1). It was disputed by Fairbairn and a number of later 'Independents' in the UK and by neo-Freudians (Sullivan, Horney, Fromm) in the USA in the 1930s and 1940s. However, these internal critics had offered rival accounts of the structure and development of the mind. One novel response to the Hook symposium, not only by psychoanalysts but also by sympathetically minded philosophers, such as Ricœur (1970) and Habermas (1972), was to question the truth and usefulness of *any* such theory within psychology.

Under pressure from these criticisms, by 1976, the first steps towards restricting psychoanalysis to clinical phenomena had been taken (Holt 1976; Klein 1976; Rubinstein 1976; Schafer 1976). It was argued that the metapsychology and associated etiological claims, e.g. the theory of libido and its development or of the superego and its origin in Oedipal conflict, should be jettisoned, as these parts of psychoanalytic theory are most difficult to derive from the clinical data and are tangential to the therapeutic relationship. From this point onwards, the scope of the term 'psychoanalysis' within the debate became unclear—was the writer concerning themselves with just the clinical relationship and the understanding of unique individuals, or a general theory of psychological functioning, and if the latter, was that theory experience-near or abstract?

Rejecting metapsychological and etiological claims was widely thought to have implications for thinking that psychoanalysis is a science. If psychoanalysis were to be a natural science, it would be so in virtue of something like Freud's economic and drive theories, for it is here that psychoanalytic theory touches most closely upon the established natural sciences of biology and neurology (see also Hopkins, this volume). Hence, a rejection of the metapsychology was also widely understood as a rejection of the idea of psychoanalysis as a science in favour of an alternative model of psychoanalytic knowledge. Making this move, it was thought, means that the critique of psychoanalysis as unscientific now misses its mark (Ricœur 1970). For psychoanalysis to fail to meet the

evidential standards of (natural) science is no failing at all if psychoanalysis is not, and does not aspire to be, a science.

The alternative that was taken up came, in the first instance, from 'hermeneutics'—the interpretative understanding of meaning. For instance, according to Habermas (1972), psychoanalytic psychotherapy is a process of (assisted) self-reflection, self-interpretation, and self-formation. This process can be transformative by opening up new ways of understanding ourselves and the meanings of our behaviour, providing insights into previously unrecognized motives. The theoretical framework of psychoanalysis—insofar as it is retained at all—must therefore be an interpretation of this process in (p. 412) general terms. Many in the debate assimilated 'science' to 'natural science', and assumed that science deals in objective causal processes. Hence psychoanalysis is not science, as it deals in meanings, not causes, and these meanings are subjective, not objective.

By the 1980s, this line of thought raised questions about the 'truth' of interpretations and the corresponding ontology of the mind (Mitchell 1988; Spence 1982). Arguments appeared claiming that the study of subjectivity—which psychoanalysis undoubtedly is—could only be a subjective study. But it is unfortunate that such associations were made so quickly and widely, as claiming that psychoanalysis is a hermeneutic discipline does not require one to accept that it is not a science. To arrive at this conclusion, one must first argue that meanings cannot be studied using a scientific method, and to defend that claim, one must have some idea both of the nature of meanings and the range of methods that count as 'scientific'.[2]

The Analytic Relationship

Central to the development of the debate in the 1980s and 1990s was a revised understanding of the relationship between analyst and analysand. Freud's scientific understanding of psychoanalysis was mirrored, it was claimed, in the tenets and practices of classical Freudian analysis (e.g. Mitchell 1998). The analyst is akin to the scientific psychologist, in that psychoanalysis seeks to discover 'the truth' about the analysand's unconscious mind, and the same truth—being part of the individual analysand's mind—ought to emerge irrespective of the particular identity of the analyst. The analyst retains 'neutrality', in part so as not to influence the clinical data produced by the analysand in ways that are tantamount to suggestion. The discovery of the analysand's unconscious states, in the right conditions—'insight'—will materially contribute to the therapeutic action of psychoanalysis.

All this came under close scrutiny, and a new recognition arose of the importance and contribution of the individual subjectivity of the analyst to the analytic relationship (Gill 1983, 1984; Hoffman 1983). It became widely accepted that the analyst is not a 'blank screen' onto which the analysand projects in the transference, as Freud thought. But Gill and Hoffman went further to argue for a new understanding of transference, a 'social conception' of transference. First, transference is, in part, a response by the analysand to the actual analyst, not only the projected one. Second, 'interpersonal reality' has different

interpretations that may be valid—it can no longer be assumed that the analysand's
(p. 413) perspective on the analytic relationship is distorted by transference while the
successfully analysed analyst's perspective is not (in theory at least). There can be no
simple division of interpersonal experience into veridical (how things really are) and
distorted. The analyst has biases and a subjective history that influence his or her
experience of the analytic relationship and the resulting interpretations are unavoidably
'subjective', in the sense of being a reflection of the subjectivity of the analyst. And so,
third, objective neutrality is impossible and, it was argued, clinically unhelpful.

This new conceptualization of the analytic relationship led simultaneously in two
directions—towards both a new epistemology and a new ontology of the unconscious.
Unfortunately, as we will see, the issues of epistemology and ontology weren't always
kept clear. Nor was the scope of 'the unconscious' that was being reconceptualized—e.g.
was the debate about unconscious experience in the clinical setting, or did it extend to
include particular historically formed behavioural dispositions and other psychological
'structures' of particular analysands identified in the clinical setting, or did it extend still
further to include such dispositions and structures as discussed in general and abstract
theorizing?

We will turn to these developments since the 1990s in just a moment, but to make sense
of them, we must first note that the developments were mediated by another theoretical
'innovation'. Having rediscovered countertransference, American analysts rediscovered
object relations (from the perspective of many British psychoanalysts, for whom such
claims were already entrenched by the 1950s, this was all at least thirty years behind, but
better late than never). One of the leading thinkers in this development, whose work shall
form the primary focus of our discussion, was Stephen Mitchell, who stated (1988: 17):
'*Mind has been redefined from a set of predetermined structures emerging from inside
an individual organism to transactional patterns and internal structures derived from an
interactive, interpersonal field*'. Rather than mental states—or at least unconscious
mental states—being expressions of endogenous forces (drives), they are *essentially*
interpersonal. '[T]he interpersonal and intrapsychic realms create, interpenetrate, and
transform each other' (Mitchell 1988: 9). The psychoanalytic school of 'relationalism'
began.

On its own, this adoption of object-relations theory is independent of Gill and Hoffman's
claims concerning transference and the analyst's knowledge, let alone the developments
that followed—as evidenced by the history of object-relations theory and transference in
Britain (Hughes, this volume). It is perfectly possible for an object-relations theorist (e.g.
Rustin 1997; Eagle 2011) to hold that there are facts about an analysand's object
relations and that one task of the analyst is to discover, with the analysand, what these
are (a second, intertwined task being to transform them). Given that the analysand
formed these object relations via their experience prior to beginning the analytic
relationship, it is even possible that the same discoveries would be made independent of
who the analyst is: 'that it may be impossible to develop any mental states … without

social interaction, does not alter the fact that, once formed, they do exist independently of us and of our "interpretive construction" '(Eagle, Wolitzky, and Wakefield 2001: 469).

(p. 414) But with this last phrase—'interpretive construction'—we are getting ahead of ourselves. Suffice to say that the development of object-relations theory from the 1980s in the USA *was* informed by the revisionary understanding of the analytic relationship presented by Gill and Hoffman, and further developed by Mitchell (1988), Renik (1993), Stolorow and Atwood (1992), Orange (1995), Aron (1996), Ogden (1997), Hoffman (1998), and others. These authors developed and defended the ontological and epistemological claims of constructivism described in the opening paragraph of this chapter, that psychological structures are built from patterns of interpersonal interaction and that individuals can only be understood within an 'interactional field'.

The Confusion Over 'Constructivism'

A first obvious epistemological implication of these developments, rightly drawn and widely granted, is fallibilism. In classical Freudian practice, it was said, analysts had been too confident in their declarations and authoritarian in their relations with patients. Greater caution and modesty were called for.

But debate was sparked by the defence of 'constructivist' readings of psychoanalytic knowledge. In the first instance, what is 'constructed', on these readings, is the *understanding* of the analysand's material that is reached by analyst and analysand together through their interpretations of their experiences of the interpersonal relationship. Reaching this understanding is not simply a matter of inferring, from the evidence, the analysand's unconscious mental states as they are already and in any case. Rather, that this understanding is *constructed* should be understood to mean that the meaning of this material is not 'there' in the mind of the analysand in the form of unconscious mental states with determinate intentional content waiting to be discovered. Compare: the realist holds that there is some definite truth about someone's mental states, some 'real' meaning to what they do, think, and feel, but to discover this, we will often need to infer those mental states from the behavioural data. The fallibilist realist accepts, further, that the evidence often supports more than one inference, and that different interpretations of others' behaviour may legitimately emphasize different mental states. But these concessions on the indeterminacy of evidence are compatible with the thought that the meaning of one's thoughts, feelings, and behaviour is given prior to and independently of the interactive relationship with another person, and thus can be said to be discovered. The constructivist rejects this model.

Beyond this, the further commitments of constructivism, its exact meaning and scope, became a matter of considerable confusion, and early, more radical statements of the position were later modified and moderated. (As noted at the outset of this chapter, I use the term 'constructivism' *not* to identify just the views of those who self-identify as

constructivists (e.g. Hoffman, Stern), but much more broadly, including other relationalists and intersubjectivists.)

(p. 415) One cause of confusion was the failure to distinguish different strengths of the epistemological claim, while a second was disagreement over the ontological implications of this claim. The starting point was, as above, that:

(1) There can be no singular, canonical interpretation of the analysand's material, as what emerges in the analytic relationship is inevitably equally a product of the analyst's subjectivity.

This claim was thought to support the further claim that:

(2) There can be no objective interpretation of the analysand's material (e.g. Renik 1993).

However, whether this is so depends on one's understanding of objectivity and its requirements (Cavell 1998). If objective knowledge requires the knower to have no subjectivity—to occupy a 'view from nowhere'—then there is no such thing as objectivity for human beings. But in that case, there is no contrast to be drawn between, say, psychoanalysis and physics. Clearly, though, claims concerning the involvement of the analyst's subjectivity are intended to draw a *contrast* with how a physicist's subjectivity impacts on physics. We will return to this in 'Beyond Objectivism and Subjectivism?'

(1) and (2) were taken together to support the stronger claims that:

(3) Analysts cannot approach the 'real' meaning of the patient's experiences (Mitchell 1998), nor should they seek to do so, as

(4) The very idea of an 'objective psychological reality' that is independent of a particular perspective on it threatens a return to the 'epistemological authoritarianism' that characterized classical Freudianism (Hoffman 1999).

From such claims concerning epistemology and clinical methodology, many authors drew a further conclusion (while some did not, and many others were simply silent on the issue):

(5) There *is* no correspondence in the patient's material to a single interpretation.

It is not simply that no one interpretation is better justified than others; there is no 'real' meaning, in that neither the patient's material nor their mental states are structured in a way that *could* correspond to one interpretation. Interpretations cannot be 'true' or 'false' in the sense of corresponding to something that exists independently of the interpretation (Mitchell 1998).

In accounting for why this is so, constructivists crossed from epistemology to ontology. On Freud's realist view, interpretation describes—accurately or not—the unconscious mental states of the analysand, which exist and have their determinate intentional content independent of the interpretation. The function of language here is descriptive.

By contrast, on constructivist views, what is unconscious does not pre-exist interpretation in a determinate way. The function of language in interpretation is creative—it brings into being determinate mental states where none existed. Thus Stern (1997: 5) talks of the analysand's *experience* itself being constructed through interpretation, and not merely the understanding of that experience. At least in part, the psychological reality of (p. 416) the analysand is constituted in and through the interaction with the analyst and the language they both use.

I say 'at least in part' as the debate quickly focused on the scope of the claim. There were disagreements, as well as changes of mind, among those who self-consciously rejected the 'realism' label, as well as misunderstandings by those who adopted it. Is all of human psychology, or just that which is relevant to psychoanalysis, to be understood thus—as a construction out of interaction? And is it a construction of present interaction or past interaction, or an interaction of present and past interaction? Is the claim restricted to manifest clinical phenomena or does it cover the objects of psychoanalytic theory as well?

In developing and defending the five constructivist claims and their ontological implications, psychoanalysts appealed to innovations in philosophy and culture concerning the status of science and the nature of human knowledge. Thus the question of realism versus constructivism concerning the unconscious developed into a debate about objectivism in human knowledge, the influence of Kuhn and Feyerabend in the philosophy of science, whether constructivism 'descends' into postmodernist relativism, and the proper interpretation of Gadamer, among other matters. This did nothing to clarify the debate (Mills 2005), which by the late 1990s became marked by repeated accusations of oversimplifications, misinterpretations (wilful as well as unwitting), obfuscation, and increasing ill humour.[3]

Narrowing the Gap

I have detailed the development of constructivism. But I also noted Eagle's comment that *Freud's* realism has largely been left behind; just as constructivists have modified their views over time, so too have realists. As object relations are central to the family of constructivist theories discussed so far, it is worth briefly outlining an understanding of object-relations theory in contemporary realist terms that eschew Freud's metapsychology.

We can understand 'object relations' as 'the cognitive, affective and motivational processes mediating interpersonal functioning, and … the enduring patterns of interpersonal behavior that draw upon these intrapsychic structures and processes' (Westen 1990: 686). Internalized 'objects', in this sense, include 'schemas' or 'models' of oneself, other persons, particular relationships, and relationship structures. These schemas are set up in childhood, derived from repeated experiences of interpersonal relating. (p. 417) Current interpersonal relationships are impacted upon by these enduring patterns, assimilating our present to our past: when some element of the schema is 'activated', e.g. experiencing a particular interpersonal situation in a certain

way, other elements of the schema come into play to interpret experience further in accordance with the schema and to influence behaviour. Hence, unconscious models of ourselves and significant others, deriving from the past, affect our perception, cognition, motivation, and affective states in myriad, patterned ways. For example, psychological defences are, to some significant extent, the result of patterns of experiencing or relating to others that derive from one's individual history of psychological development. As the result of both developmental factors (the deficit model of psychopathology) and the motivation to avoid psychological pain (the defence model), these models may inaccurately reflect social and intrapsychic reality. Or again, this process underpins the processes involved in transference. A minimal resemblance of the analyst to an analysand's 'significant object' leads to a set of responses, attitudes, and behaviours deriving from the patient's object relation (Andersen and Thorpe 2009).

On a realist reading, object relations are understood as psychological structures that, through interaction with real-time interpersonal situations, causally explain a subject's occurrent conscious and unconscious psychological processes and states. Object relations are understood as mediating the causal effects of past experience in which they originate on both present dispositions and, through these, on present experience. In this sense, it is indeed true to say that the intersubjective psychological reality of the consulting room is a construction from the subjectivity of both analysand and analyst.[4]

How do constructivists respond to this picture? Some critics of constructivism (e.g. Meissner 1998: 422) have understood constructivists to claim that all mental states of an analysand are 'constructed' in the interpersonal relationship with the analyst, such that the analysand does not bring with him or her any dispositions—most importantly, object relations—that existed prior to encountering the analytic situation. But this is a misunderstanding. Stolorow and Atwood (1992: 24) clarify:

> Some may see a contradiction between the concept of developmentally preestablished principles that organize subsequent experiences and our repeated contention that experience is always embedded in a constitutive intersubjective context. This contradiction is more apparent than real. A person enters any situation with an established set of ordering principles (the subject's contribution to the intersubjective system), but it is the context that determines which among the array of these (p. 418) principles will be called on to organize the experience. Experience becomes organized by a particular invariant principle only when there is a situation that lends itself to being so organized. The organization of experience can therefore be seen as codetermined *both* by preexisting principles *and* by an ongoing context that favors one or another of them over the others.

Similarly, Stern (1997: 42) acknowledges 'character, schemata, transference … internal object representations … unconscious fantasies … the plethora of processes in any one person's psychic life that keep reproducing experience in meaningfully similar shapes', and says that childhood experiences 'structure … the possibility of knowing—the potential for what we can say and think and what we cannot' (Stern 1997: 31). And Mitchell

(2000a: 155) comments that 'There is no place in any of my writings in which I argue against the idea that the patient has a mind with preexisting properties before ever encountering the analyst or that there are no continuities among the versions of ourselves that emerge with different people. This would, of course, be preposterous.'[5] Finally, Benjamin (1999: 886) notes the shift that has occurred in Hoffman's constructivism. While in the 1980s Hoffman spoke of 'social' constructivism to emphasize the interpersonal nature of the process, by the late 1990s, he talks of 'dialectical' constructivism, in which the constructive process is in dialectical exchange with what is already 'given' for the individual. Like Mitchell, Hoffman gives explicit recognition to the history that the individual brings to any interpersonal relationship.

What, then, is the disagreement between revised realism and revised constructivism? Benjamin (1999: 887) suggests that:

> the issue that divides us is not the existence of individual history but the question of how we come to know and interact with the patient's history, fantasies, and conflicts—via our own subjective experience or through objective knowledge—and whether these are seen as ambiguous but not absolutely malleable or amorphous.

Here again we see both the epistemological claim and the ontological claim concerning determinacy. A *kind* of realism has been granted, in the form of psychological structures—of whatever kind—that carry an individual's psychological history into the present. But because constructivists' primary target has been *Freud's* realism, with its claim that analysis discovers the analysand's unconscious mental states as they exist independently of interaction with the analyst, constructivists have often felt that properly recognizing the influence of the situation, especially the analyst's subjectivity, on the (p. 419) experience of the analysand and its interpretation is inconsistent with realism. Thus Mitchell (2000a: 155) continues, 'a problem with preconstructivist thinking is the assumption that there is a static organization to mind that manifests itself whole cloth across experiences'. He goes on to approve of Ogden's (1997: 190) description of the interaction between object relations and situations:

> The internal object relationship … is not a fixed entity; it is a fluid set of thoughts, feelings, and sensations that is continually in movement and is always susceptible to being shaped and restructured as it is newly experienced in the context of each new unconscious intersubjective relationship. In every instance it will be a different facet of the complex movement of feeling constituting an internal object relationship that will be most alive in the new unconscious intersubjective context.

This last sentence repeats Stolorow' and Atwood's claim that which element of an object relation—which 'principle', which 'facet'—is activated in a given situation, and the consequent effects on experience and behaviour, are importantly a function of the situation.

But while this interactive understanding may conflict with some, especially early, object-relations theories, it is perfectly compatible with more recent theories as they are outlined earlier. To some extent, so too is Ogden's claim that an object relation is 'always susceptible to being shaped and restructured as it is newly experienced in the context of each new unconscious intersubjective relationship'. While realist object-relations theories typically understand object relations as stable, enduring structures, this is compatible with the thought that they are *susceptible* to change through new experiences. Realists in the debate, then, have quickly conceded that Freud's realism, with its thought that what emerges in analysis are the unconscious states as they pre-existed in the analysand's mind, is untenable. They accept that a much more dynamic, interactive understanding of the unconscious is necessary. But, they argue, this needn't lead us to abandon realism per se (Eagle, Wolitzky, and Wakefield 2001: 477).

The Remaining Debate

Is that it, then? Is the debate now resolved? Unfortunately not. Questions remain concerning how determinate object relations are and whether there is any meaningful sense in which a psychoanalyst's knowledge may be said to be 'objective'. I suggest, oversimplifying somewhat, that the debates over both questions are fed by the two sides focusing on different aspects of 'psychoanalysis', identified in the Introduction. The constructivists are primarily concerned with individual analysts' *knowledge of individual analysands*, and tend to reject or remain silent on the possibility of a general empirical theory of the mind. The realists, while also concerned with the clinical encounter, are equally concerned with the development of more *abstract theoretical knowledge*. Let me explain.

(p. 420) In 'The Critique of Freud', I noted that constructivism has its origins in a rejection of Freudian metapsychology and the restriction of psychoanalysis to clinical phenomena. The clinical encounter provides the experience that defines the object of psychoanalytic investigation. Thus, the primary concern of constructivists, and the primary term in which to think about the ontology of mind, is experience. And experience is what is said to be constructed within an interpersonal interaction. Experience, of course, is *occurrent*, not dispositional, and the construction of experience is a complex *process*—hence 'mind' is conceptualized as unfolding in the present. But experience is not determinate. Psychoanalysis makes sense of experience through identifying some patterns among others that may justly be found in the occurrent psychological phenomena, past and present.

The unconscious in psychoanalysis, on this approach, is to be defined in terms of the experience of the clinical encounter—it is what is unconscious in this unique experience of relationship. Thus, the unconscious is first and foremost defined in terms of what is outside the awareness of the subject yet informs and shapes their intersubjective experience. In formulations that are reminiscent of Merleau-Ponty (see Phillips, this volume, and Fuchs, this volume), Stern (1997) conceives of this as part of the subject's

experience, but left 'unformulated' rather than brought to consciousness, while Stolorow (2001: xii) conceives of it as falling 'outside the horizons of a person's experiential world'. Insofar as there can be a general account of the mind, it is hermeneutic and phenomenological rather than taking the form of a 'scientific theory'. Furthermore, what is unconscious for a subject is a function not of the individual subject, as realist object-relations theory is sometimes still prone to theorizing it, but of the unique relationship between two subjects, the analysand and analyst. Thus, we arrive at the claim that the concept of the 'intersubjective system' (Orange, Atwood, and Stolorow 1997: 76), the 'analytic third' (Ogden 1994), or the 'interpersonal field' (variously conceptualized) is primary in clinical work.[6]

From the constructivist perspective, the way realists are prone to understand object relations, even after acknowledging their indeterminacy, is a misleading theoretical hypostatization. Psychoanalysis, constructivists claim, is not seeking to identify what realists think of as existing mental states as they are already (as though 'fixed' or 'static'). Hence the claim that there is no one 'correct' interpretation of experience, no 'given' meaning.[7] Rather, *what* is known is indeterminate and *how* it is known is through (p. 421) co-construction. 'Psychoanalysis' refers to the clinical encounter and 'psychoanalytic knowledge' is the knowledge individual analysts have of individual analysands.

Realists are happy to grant that psychoanalytic knowledge begins in the clinical encounter, but do not limit it to this. We have seen that realists insist on the importance of object relations as ways in which the mind of an analysand is 'preorganized' in advance of the clinical relationship. They appeal to clinical experiences of 'discovery' and of how an interpretation can seem to 'fit' what is 'already' part of the analysand's unconscious. But constructivism will claim to explain such experiences equally well. First, such experiences take place within the co-constructed clinical encounter. And within this unique clinical relationship, discoveries may be made and interpretations may well 'fit'. But this does not demonstrate realism, since the phenomenology is (partly) a product of the relationship. Second, with revised constructivism, the claim is not that there is no psychic reality to be encountered, but that such reality is indeterminate and that knowledge of it cannot be objective for the reasons rehearsed earlier.

However, realists hold that psychoanalysis is not only about understanding individuals within a unique relationship. It is also the construction of a general psychological theory, based on information about the *generic* processes and obstacles present in such interpersonal experiences. Claims about transference, the many patterns of psychic defence, resistance, conflict and compromise, the understanding of dreams, unconscious emotions, the symbolic content of motives, and the influence of the past are part of this psychological theory. This theoretical apparatus, and the concepts it employs, enables us to understand, explain, and in some cases predict human behaviour. This conceptual and theoretical structuring and organization of our ever-changing conscious and unconscious experience is what gives realism its motivation and grip in seeking to understand and explain why experience takes the structure it does.

Clinical practice, says the realist, requires such a more general theoretical backdrop, and so *any* psychoanalytic approach to understanding ourselves, including constructivism, must take some stance on these theoretical claims. Indeed, these general claims concerning the clinical process are inescapable. Their general and theoretical status is hidden in constructivism either because they are silently assumed or because they are thought to be a matter of phenomenological hermeneutics and thus not 'theoretical'. But they nevertheless inform the process of making sense of experience. Constructivists continue to approach the task of interpreting the individual within the framework of a psychoanalytic *model* of the mind. But, realists point out, this model has been disputed by other psychological theories, and can therefore legitimately be understood to comprise an alternative theory which can, in an important sense, be thought to be true or false.[8]

(p. 422) These differences in what is meant by 'psychoanalytic knowledge' feed the debate over its possible objectivity. And thus we see the influence of restricting the meaning of 'psychoanalysis' to the clinical encounter or extending it to include a psychological theory, creating a tendency for those on each side of the debate to misunderstand the scope and approach of the opposition.[9]

Continuing the realist line of thought a moment longer, it may seem that if it is true that our experience is inflected by transference, defence, and the like, then what makes this true are determinate mental operations on determinate mental contents. If there really are such processes, and there really are such contents, then this secures a particular description of a person as, say, projecting their fear, as objectively true. But does this follow? It is claims like these that constructivists want to challenge most strongly. Is talk of 'objectivity' or even 'truth' appropriate here? Would some other *telos*, such as understanding, usefulness, or praxis, provide a better model? These are concerns that constructivism presses, appealing to the reconsideration of objectivity and truth in philosophical debates concerning the status of science, within both 'analytic' philosophy of science following Kuhn (1970) and hermeneutic theories, perhaps most especially that of Gadamer (1975).

Beyond Objectivism and Subjectivism?

At this point, we must leave the general debate to continue with a single point of focus. I shall consider the issue through the lens of Mitchell (2000a). He says that 'the major point' of his substantial (1998) article was to show that the oppositional categories of 'objective' and 'subjective' are no longer helpful. Instead, he claims, recent debates in psychoanalysis and philosophy point to a third category that is neither objective nor subjective, as exemplified in Bernstein's (1983) *Beyond Objectivism and Relativism*. I am fundamentally in agreement with Bernstein, but I shall argue that Mitchell has misunderstood him.

Bernstein (1983: 74) argues that the upshot of debates in philosophy of science is 'an appreciation of the practical character of rationality in science'. This practical character includes 'the role of choice, deliberation, conflicting variable opinions, and the

judgmental quality of rationality'. As a result, we can no longer think of 'the' scientific method (is there just one?) as 'a set of permanent, unambiguous rules' nor can we assimilate scientific argument to 'models of deductive proof or inductive generalization'. Instead, to understand 'the rationality of scientific inquiry we must focus on the conflict of theories, paradigms, research programs, and research traditions in their *historical* (p. 423) *development*' (see Lakatos 1978) and 'the nature, function, and dynamics of *communities of inquirers*' (Bernstein 1983: 77).

This phrase—community of enquiry—comes from American pragmatist Charles Peirce. He argues that 'the conception of reality … essentially involves the notion of a COMMUNITY, without definite limits, and capable of a definite increase of knowledge' (1931–5: Vol. 5, para. 311, 186–7). While Bernstein and Peirce are primarily concerned with science, we can extend the point to all rational enquiry. Our account of the nature, function, and dynamics of such communities of enquiry must include the recognition of the *practical norms* which they take as constitutive of their enquiry and intersubjective communication, i.e. the norms that govern what counts as giving a reason, when evidence is sufficient, how claims are to be defended or critiqued, and so on.

Thus, says Bernstein, we must give up 'objectivism' *understood as* the claim that 'there is or must be some permanent, ahistorical matrix or framework to which we can ultimately appeal in determining the nature of rationality, knowledge, truth, reality, goodness, or rightness' (1983: 8). Instead, rationality is dialogic, intersubjective, and historically situated—these are all ways in which reasoning, even scientific reasoning, is practical. However, this does not yield relativism or subjectivism. Giving up objectivism as understood, Bernstein argues, is perfectly compatible with recognizing that 'in the course of scientific development the cumulative weight of evidence, data, reasons and arguments can be rationally decisive for scientific communities' (1983: 74). Kuhn, Lakatos, and Peirce all acknowledge that science progresses and that this is an achievement of rationality. Recognizing the practical grounding of rationality and the failure of *objectivism* is no attack on *objectivity* understood thus (Bernstein 1983: 92).

Bernstein (1983: 154) goes on to argue that Gadamer similarly understands truth as 'what can be argumentatively validated by the community of interpreters who open themselves to what tradition "says to us" '. Fundamental to Gadamer's epistemology is the concept of 'dialogue', 'a process of two people understanding each other' (Gadamer 1975: 347). This requires an openness to the other and to new experience (1975: 319), which can never be completed. While this makes claims to truth fallible and open to criticism, Bernstein argues, it doesn't remove the need for 'validation that can be realized only through offering the best reasons and arguments that can be given in support of them—reasons and arguments that are themselves embedded in the practices that have been developed in the course of history' (1983: 168). Bernstein furthermore points out that Gadamer fails to recognize the *general* application of these ideas, frequently—but misleadingly—speaking of hermeneutical understanding as 'an "entirely different type of knowledge and truth" from that which is yielded by … science' (1983: 168). We should correct this in light of the earlier conclusions concerning science and rationality. Within

science, the scientific method provides a recognized form of validation, a manifestation of the demand for intersubjective criteria of truth within a community of enquirers.[10]

(p. 424) How is this revised understanding of science, rationality, and truth supposed to be helpful to the constructivist defence of psychoanalytic knowledge? Silverman (2000: 148) suggests that Mitchell (1998) should be understood as defending 'relative objectivity'. In light of the outline of Bernstein's argument, and Mitchell's apparent endorsement of it, this looks like a reasonable suggestion. However, Mitchell (2000a) explicitly rejects this, arguing that not only are the old dichotomies of 'objective' and 'subjective' unhelpful, but that combining them as Silverman suggests is 'obfuscating'. To justify his rejection of Silverman's suggestion, Mitchell says:

> psychoanalytic knowledge is not usefully considered objective in any traditional sense of that word. The analyst can never be confident that he or she can, through self-reflection, filter out his or her own subjectivity, through which he or she has brought to life and come to understand the patient. But it is also not helpful to consider psychoanalytic knowledge relativistic. Good analysts do not simply make things up. Good analysis is highly disciplined, careful, thoughtful, continually self-reflective, and responsive to subsequent experience … [B]ecause neither objectivism nor relativism seems an adequate description of psychoanalytic knowledge, combining them oxymoronically just compounds the problem … It does not help to imply that we can come close, or sort of, or almost, to separating out the analyst's involvement. (2000: 157–8)

However, these grounds for rejecting 'objectivity' have almost nothing at all to do with Bernstein's understanding of knowledge, truth, and reason in terms of a third category 'beyond objectivism and relativism'. Mitchell's appeal to Bernstein misses its mark for two reasons.

First, Mitchell confuses objectivism and objectivity. Bernstein's arguments apply equally to physics, chemistry, and biology, to palaeontology, economics, and history. Thus, his arguments against objectivism leave completely open whether psychoanalytic knowledge has the same structure and grounding as the natural sciences. If Bernstein's reading of the debate in the philosophy of science is correct (and I think that, on the whole, it is), other appeals to post-Kuhnian conceptions of science and rationality in support of attacks on objectivity in psychoanalytic knowledge are equally misguided. (p. 425) Awareness of general developments in epistemology and the philosophy of science are helpful to ward off 'objectivism' as Bernstein uses the term and any resulting simplistic scientism. As long as one rejects objectivism—the idea of a permanent ahistorical set of principles for determining rationality and truth—as do the realists Cavell (1998) and Silverman (2000), there is no conflict between Bernstein's position and practice-specific conceptions of realism and objectivity.

Second, as the earlier quotation from Mitchell indicates, his primary concern is with the *individual* analyst's knowledge of the individual analysand. By contrast, Bernstein's arguments are not primarily concerned with the individual subjectivity or involvement of

the enquirer, but operate at the level of *communities* with accepted norms of reasoning and evidence. He is primarily concerned with the general or theoretical knowledge held by a discipline. The concern is with the norms around method and communication, rather than the construction of the interpersonal field within each unique analyst–analysand dyad. Once again, there has been confusion concerning what is at issue in talk of 'psychoanalytic knowledge'. It is perfectly reasonable—and I think, correct—to hold that an individual analyst's knowledge of an individual analysand cannot be 'objective' in the same sense as psychoanalytic knowledge at the level of general theory. Specific interpretations concerning the meaning of an individual's thoughts, feelings, and behaviour may not have the same epistemological status as widely accepted general claims about psychological functioning, as they are closely tied to—and influenced by— individual situations and individual enquirers. This is true not just in psychoanalysis, but in psychology as a whole (Wilson and Brekke 1994: 121). This understanding of the relationship between individual and general knowledge supports the realist reading of psychoanalytic knowledge. Once corroboration by the community of enquirers has established a general claim, e.g. concerning the occurrence of transference or a typical pattern of defence, then it can be made use of when seeking to understand meanings on any given occasion (see Lacewing 2013). It is a realist defence of objective knowledge, rather than a constructivist attack on it, that coheres best with Bernstein's analysis of the philosophy of science.

Unconscious Meaning

We are left with the remaining question of the ontology of the unconscious, and in particular, the determinacy of object relations: those unconscious structures, states, or processes that are acknowledged by both sides of the debate to be an individual's contribution to an interpersonal situation.

I begin by acknowledging again the constructivist point, made in 'The Remaining Debate', that what is unconscious for an individual is not best understood in terms of their mind taken on its own. First, what they are unconscious of in themselves is itself a product of interpersonal situations. Second, what is unconscious at any one time is a function of the intersubjective situation they are in, and thus may alter from one (p. 426) situation to another. Third, there will be features of the other and the situation, as well as of themselves, of which they will be unconscious.

Yet even if we grant that the unconscious in psychoanalysis is more about patterns of interpersonal experience than the Freudian conception of 'hidden depths' in the individual mind, are there nevertheless determinate truths about a subject's object relations—about what they do and do not, or can and cannot, experience in their relationships with others? In particular, can unconscious psychological phenomena have determinate intentional content and meaning while unconscious, or do they only acquire it upon being brought to consciousness through interpretation? My focus for this last, brief discussion will be Stern (1997).

How Should We Understand the Psychoanalytic Unconscious?

In 'The Critique of Freud' we saw that Freud's realism holds that unconscious mental states and their intentional content exist prior to and independent of their interpretation in consciousness. According to constructivism, this is not so. Starting from this thought, Stern understands the unconscious as 'unformulated experience':

> if language is not merely a set of tags or labels for experience, but actually plays a role in constituting it, we are challenged to change our conception of what it means for experience to be unconscious. Unconscious experience ... is no longer merely hidden, awaiting only language to bring it out of the shadows. Instead, the form it will eventually take in words is not predetermined by its own structure ... unconscious experience *exists*; but it does not exist in forms in which we can grasp it in words. It remains to be interpreted. Unconscious experience and meaning is what I call unformulated experience. (1998: xi)

Although Stern talks here of unconscious *experience*, it is clear from later discussions in the book that he intends these claims to apply to unconscious phenomena generally, including, for example, object relations—which he himself identifies as dispositional structures rather than experiences. Thus, he holds that unconscious contents are indeterminate and non-verbal ('unformulated'), and indeed, that they are indeterminate in some sense *because* they are non-verbal. They only acquire greater determinacy as their meaning is 'formulated', i.e. interpreted in consciousness.

But these remarks, even if true, are not sufficient to account for the contents being unconscious. There is no easy equation between what is unconscious and unformulated experience as described. There are conscious but indeterminate or unformulated experiences and states. And there are semantic unconscious experiences and states.

1. The indeterminate conscious: Much, if not all, intentional content is indeterminate to some extent (Mackie 1975). Suppose I want a cup of tea, and I take my tea with milk and sugar. If someone makes me black tea, I didn't want *that*. My desire is more determinate than wanting just anything that can be described as 'a cup of tea'. But it is not so determinate that I can be said to want it with 17 per cent milk and 78 grains of sugar. Yet any particular cup of tea will have a determinate percentage of milk and number of grains of sugar. Many cups of tea that satisfy my desire differ in these determinate respects. So the content of my desire is indeterminate, yet perfectly conscious.

(p. 427) **2.** The unformulated conscious: Non-verbal, and more generally non-conceptual, experience is unformulated but can be conscious. I look at an abstract work of art. I see the shapes and the colours. I do not (and probably cannot) put into words much about those shapes and colours. But there is nothing unconscious about my visual experience of seeing them.

3. The semantic unconscious: While controversial, there is evidence of semantic priming, i.e. subliminal processing of the meanings of words, presented either visually or aurally (e.g. Naccache and Dehaene 2001). Given that Stern understands putting something into words as making it (more) determinate, I take the occurrence

of unconscious states and processes with semantic content to demonstrate that some unconscious psychological phenomena are (at least somewhat) determinate.

A first step towards a more complete account of the relation between what is unconscious and what is unformulated involves recognizing that the psychoanalytic unconscious does not exhaust the psychological unconscious. The psychoanalytic unconscious as unformulated experience concerns particular forms of experience—experience of the self, others, and interpersonal relations. So how are we to account for both the unconscious and the indeterminate status of these?[11]

Stern assimilates unformulated experience to implicit know-how. The 'felt meaning' that is unformulated cannot be verbal or semantic meaning, nor even a condensation of these. The difficulty of finding the right words is not the result of a complexity of unconscious associations and primary processes that are in themselves each determinate (1997: 49). Rather, the meaning derives from a 'semiotics' of practice: 'Reflect on the difficulty of describing in words the details of how you ride a bicycle … The difficulty is due to the fact that these activities are not represented in words at all, but in the semiotic of practice. The same is true for the difficulty of describing how one deals with one's mother … ' (1997: 18).

But again, this is not quite sufficient. Stern's model here is missing a psychodynamic element. The way in which—or again, the reasons why—implicit know-how is unconscious can be entirely distinct from anything that psychoanalysis is interested in, entirely a matter of non-verbal procedural knowledge. And even when one turns from how to ride a bicycle to interpersonal relations, the same can hold, e.g. my unconscious knowledge of how to engage in conversation with others is implicit but not for psychodynamic reasons. And as Leite (this volume) points out, even the 'implicit relational (p. 428) knowing' (Lyons-Ruth et al. 1998) involved in intimate relationships does not yet provide a full account of the unconscious that is of interest to psychoanalysis.

However, we may provide the beginnings of an integrative model using the 'schema'-based understanding of object relations delineated above in 'Narrowing the Gap'. On this view, the form that what is interpreted takes while unconscious may well be indeterminate, requiring or inviting a 'completion' or specification in meaning as it is brought into consciousness. *What is unconscious, at root, is a psychodynamically charged and incompletely specified procedural model*—a model that may be one among others—*for how to relate self and others, including within that how to be oneself and how to understand significant others*, which model is in constant interaction with one's current situation. What one is unconscious of may become more determinate through the process of formulation, as new associations, conceptualizations, or precise feelings are formed and other possibilities are or are not taken up. But equally, as Stern (1997: 28–9) notes, this process is 'constrained by' or answerable to what is unconscious before the process of its interpretation and specification as it becomes conscious.

This proposal, I believe, is coherent with Stern's constructivism, and yet it is a development of a contemporary realist understanding of object relations. Should we say 'unconscious contents (meanings) are not fully determinate, and only become so through interaction with an intersubjective situation, and so constructivism is true'? Or should we say 'unconscious contents are partially (though not linguistically) determinate, and contribute to intersubjective situations, and so realism is true'? Mightn't we say both?

Perhaps some may struggle with the notion that what is indeterminate can be 'real'. Of course, if our touchstone for what is real is our everyday experience of physical objects, this can seem puzzling. But this just shows that we shouldn't expect psychological reality to resemble physical reality in this respect. As already noted, there is a sense in which all intentional content is indeterminate, while non-conceptual content and procedural knowledge is linguistically indeterminate. Everyone in the debate grants that the historical inheritance of an individual shapes their present experience—that this, at least, is not itself constructed through conscious interpretation alone. Everyone also grants that such an inheritance may—and in the course of a successful analysis, will—be changed for the future by precisely such conscious interpretations (though not only these). Whether we wish to emphasize, with Ogden, the fluidity and responsiveness of object relations to new relationships or, with Eagle, the stability of object relations and the difficulty of achieving lasting change, is now a matter less for the theorist than for the practitioner.

References

Altman, N. and Davies, J. M. (2003). 'A plea for constructive dialogue'. *Journal of the American Psychoanalytic Association 51S*: 145–161.

Andersen, S. M. and Chen, S. (2002). 'The relational self: An interpersonal social-cognitive theory'. *Psychological Review 109*: 619–645.

(p. 429) Andersen, S. M. and Thorpe, J. S. (2009). 'An IF-THEN theory of personality: Significant others and the relational self'. *Journal of Research in Personality 43*: 163–170.

Aron, L. (1996). *A Meeting of Minds*. Hillsdale, NJ: Analytic Press.

Benjamin, J. (1999). 'Hoffman's *Ritual and Spontaneity*: Three reviews and a response'. *Journal of the American Psychoanalytic Association 47*(3): 883–891.

Bernstein, R. (1983). *Beyond Objectivism and Relativism*. Oxford: Blackwell.

Borch-Jacobsen, M. and Shamdasani, S. (2012). *The Freud Files*. Cambridge: Cambridge University Press.

Cassidy, J. and Shaver, P. R. (1999). *Handbook of Attachment: Theory, Research and Clinical Applications*. New York: Guilford Press.

Cavell, M. (1998). 'In response to Owen Renik's "The analyst's subjectivity and the analyst's objectivity" '. *International Journal of Psychoanalysis 79*: 1195–1202.

Chang, H. (2017). 'Pragmatist coherence as the source of truth and reality'. *Proceedings of the Aristotelian Society 117*(2): 103–122.

Eagle, M. (2011). *From Classical to Contemporary Psychoanalysis: A Critique and Integration*. New York: Routledge.

Eagle, M. N., Wakefield, J. C., and Wolitzky, D. L. (2003). 'Interpreting Mitchell's constructivism'. *Journal of the American Psychoanalytic Association 51S*: 163–178.

Eagle, M. N., Wolitzky, D. L., and Wakefield, J. C. (2001). 'The analyst's knowledge and authority: A critique of the "new view" in psychoanalysis'. *Journal of the American Psychoanalytic Association 49*: 457–489.

Fairbairn, R. (1952). *Psychoanalytic Studies of the Personality*. London: Tavistock Publications.

Fleeson, W. and Noftle, E. (2009). 'In favor of the synthetic resolution to the person–situation debate'. *Journal of Research in Personality 43*(2): 150–154.

Frank, G. (1998). 'The intersubjective school of psychoanalysis: Concerns and questions'. *Psychoanalytic Psychology 15*: 420–423.

Freud, S. (1895). 'Project for a scientific psychology'. In J. Strachey (ed. and trans.), *Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 1 (pp. 281–391). London: Hogarth Press.

Freud, S. (1923). *The Ego and the Id*. In J. Strachey (ed. and trans.), *Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 19. London: Hogarth Press.

Fromm, E. (1979) *Greatness and Limitations of Freud's Thought*. New York: Harper.

Gadamer, H.-G. (1975). *Truth and Method*, trans. and eds G. Barden and J. Cumming. New York: Seabury Press.

Gill, M. M. (1983). 'The interpersonal paradigm and the degree of the therapist's involvement'. *Contemporary Psychoanalysis 19*: 200–237.

Gill, M. M. (1984). 'Transference: A change in conception or only in emphasis?' *Psychoanalytic Inquiry 4*(3): 489–523.

Gipps, R. (2018). 'Psychoanalysis: A science of the mind?' *Philosophy, Psychiatry and Psychology 25*(2): 113–118.

Habermas, J. (1972). *Knowledge and Human Interests*. London: Heinemann Educational.

Hoffman, I. Z. (1983). 'The patient as interpreter of the analyst's experience'. *Contemporary Psychoanalysis 19*: 389–422.

Hoffman, I. Z. (1998). *Ritual and Spontaneity in the Psychoanalytic Process*. Hillsdale, NJ: Analytic Press.

Hoffman, I. Z. (1999). 'Response'. *Journal of the American Psychoanalytic Association 47*(3): 907–920.

Holt, R. R. (1976). 'Freud's theory of the primary process: present status'. *Psychoanalysis and Contemporary Science 5*: 61–99.

(p. 430) Hook, S. (ed.) (1959). *Psychoanalysis, Scientific Method and Philosophy*. New York: New York University Press.

Jaspers, K. (1913/1963). *General Psychopathology*, trans. J. Hoenig and M. W. Hamilton. Chicago, IL: University of Chicago Press.

Klein, G. S. (1976). *Psychoanalytic Theory: An Exploration of Essentials*. New York: International Universities Press.

Kohut, H. (1971). *The Analysis of the Self*. New York: International Universities Press.

Kuhn, T. (1970). *The Structure of Scientific Revolutions*, 2nd edn. Chicago, IL: University of Chicago Press.

Lacewing, M. (2013). 'The problem of suggestion: an analysis and solution'. *Philosophical Psychology 26*: 718–743.

Lacewing, M. (2018). 'Evidence, inference and causal explanation: a reply to Gipps'. *Philosophy, Psychiatry and Psychology 25*(2): 119–122.

Lakatos, I. (1978). *The Methodology of Scientific Research Programmes*. Cambridge: Cambridge University Press.

Lyons-Ruth, K. et al. (1998). 'Implicit relational knowing: Its role in development and psychoanalytic treatment'. *Infant Mental Health Journal 19*(3): 282–289.

Mackie, J. (1975). 'Problems of intentionality'. In J. Mackie, *Logic and Knowledge*. Oxford: Oxford University Press.

Meissner, W. W. (1998). 'Review of S. A. Mitchell's "Influence and autonomy in psychoanalysis" '. *Psychoanalytic Books 94*: 419–423.

Mills, J. (2005). 'A critique of relational psychoanalysis'. *Psychoanalytic Psychology 22*(2): 155–188.

Mitchell, S. (1988). *Relational Concepts in Psychoanalysis*. Cambridge, MA: Harvard University Press.

Mitchell, S. (1998). 'The analyst's knowledge and authority'. *Psychoanalytic Quarterly 67*(1): 1–31.

Mitchell, S. (2000a). 'Reply to Silverman'. *Psychoanalytic Psychology 17*: 153–159.

Mitchell, S. (2000b). *Relationality*. New York: Psychology Press.

Mitchell, S. and Aron, L. (eds) (1999). *Relational Analysis*. Hillsdale, NJ: Analytic Press.

Moran, R. (2001). *Authority and Estrangement*. Princeton, NJ: Princeton University Press.

Murphy, D. (2007). *Psychiatry in the Scientific Image*. Cambridge, MA: MIT Press.

Naccache, L. and Dehaene, S. (2001). 'Unconscious semantic priming extends to novel unseen stimuli'. *Cognition 80*(3): 215–229.

Ogden, T. (1994). 'The analytic third: Working with intersubjective clinical facts'. *International Journal of Psychoanalysis 75*: 3–19.

Ogden, T. (1997). *Reverie and Interpretation*. Northvale, NJ: Aronson.

Orange, D. (1995). *Emotional Understanding*. New York: Guilford Press.

Orange, D. M., Atwood, G. E., and Stolorow, R. D. (1997). *Working Intersubjectively*. Hillsdale, NJ: Analytic Press.

Peirce, C. S. (1931–5). *Collected Papers of Charles Sanders Peirce*, eds C. Harteshorne and P. Weiss. Cambridge, MA: Harvard University Press.

Putnam, H. (1981). *Reason, Truth and History*. Cambridge: Cambridge University Press.

Putnam, H. (1992). *Realism with a Human Face*. Cambridge, MA: Harvard University Press.

Renik, O. (1993). 'Analytic interaction: Conceptualizing technique in light of the analyst's irreducible subjectivity'. *Psychoanalytic Quarterly 62*: 553–571.

Renik, O. (1999). 'Renik replies to Cavell'. *International Journal of Psychoanalysis 80*(2): 382–383.

(p. 431) Richards, A. D. (2003). 'Psychoanalytic discourse at the turn of our century: A plea for a measure of humility'. *Journal of the American Psychoanalytic Association 51S*: 73–89.

Ricœur, P. (1970). *Freud and Philosophy*. New Haven, CT: Yale University Press.

Rubinstein, B. (1976). 'On the possibility of a strictly clinical psychoanalytic theory: an essay in the philosophy of psychoanalysis'. *Psychological Issues 9*(4): 229–264.

Rustin, M. J. (1997). 'The generation of psychoanalytic knowledge: Sociological and clinical perspectives part one: "Give me a consulting room: psychoanalysis and science studies" '. *British Journal of Psychotherapy 13*(4): 527–541.

Schafer, R. (1976). *A New Language for Psychoanalysis*. New Haven, CT: Yale University Press.

Silverman, D. K. (2000). 'An interrogation of the relational turn: A discussion with Stephen Mitchell'. *Psychoanalytic Psychology 17*: 146–152.

Spence, D. (1982). *Narrative Truth and Historical Truth*. New York: Norton.

Stern, D. B. (1997). *Unformulated Experience*. Hillsdale, NJ: Analytic Press.

Stern, D. B. (2003). 'Commentary on "Psychoanalytic discourse at the turn of our century: A plea for a measure of humility" '. *Journal of the American Psychoanalytic Association 51S*: 98–108.

Stern, D. N. (1985). *The Interpersonal World of the Infant*. New York: Basic Books.

Stolorow, R. D. (1998). 'Clarifying the intersubjective perspective: A reply to George Frank'. *Psychoanalytic Psychology 15*: 424–427.

Stolorow, R. D. (2001). 'Foreword'. In P. Buirski and P. Haglund, *Making Sense Together: The Intersubjective Approach to Psychotherapy*. New York: Jason Aronson.

Stolorow, R. and Atwood, G. (1992). *Contexts of Being: The Intersubjective Foundations of Psychological Life*. Hillsdale, NJ: Analytic Press.

Stolorow, R. and Atwood, G. (2016). 'Experiencing selfhood is not "a self" '. *International Journal of Self Psychoanalysis 11*(2): 183–187.

Stolorow, R., Atwood, G., and Orange, D. (2006). 'Contextualizing Is Not Nullifying: Reply to Mills (2005)'. *Psychoanalytic Psychology 23*(1): 184–188.

Taylor, C. (1964). *The Explanation of Behaviour*. London: Routledge & Kegan Paul.

Taylor, C. (1985). 'Self-interpreting animals'. In C. Taylor, *Human Agency and Language*. Cambridge: Cambridge University Press.

Westen, D. (1990). 'Towards a revised theory of borderline object relations: contributions of empirical research'. *The International Journal of Psychoanalysis 71*: 661–693.

Wilson, T. and Brekke, N. (1994). 'Mental contamination and mental correction: Unwanted influences on judgments and evaluations'. *Psychological Bulletin 116*(1): 117–142.

Wright, C. (1991). *Truth and Objectivity*. Cambridge, MA: Harvard University Press.

(p. 432)

Notes:

(¹) Thanks to Richard Gipps, Adam Leite, and Katy Abramson for their helpful and insightful comments. An earlier version of this chapter was presented at the conference 'Modes of the Unconscious' in Heidelberg in October 2016. Thanks to the organizers, Stefan Kristensen and Thomas Fuchs, and to participants for their questions and suggestions.

(²) That psychiatry could be both hermeneutic and a science was, of course (though rarely noted), Jaspers's (1913/1963) view, which he further articulated in terms of phenomenological investigation, and some contemporary phenomenological engagements with psychoanalysis continue to defend this scientific understanding (Fuchs, this volume). Also frequently overlooked was the possibility that psychoanalysis is a social science, and social sciences are genuine forms of science, though with appropriately different methodologies and concerns from natural sciences, given their hermeneutic nature (see Lacewing 2013).

(³) For example, see Stolorow's (1998) response to Frank's (1998) review; Renik's (1999) response to Cavell's (1998) review; Mitchell's (2000a) response to Silverman's (2000) review; Stern's (2003) response to Richards's (2003) review; Altman and Davies's (2003) response to Eagle, Wakefield, and Wolitzky's (2001) review and Eagle, Wakefield, and Wolitzky's (2003) reply in turn; Stolorow, Atwood, and Orange's (2006) response to Mills's (2005) review.

(⁴) As discussed further in 'The Remaining Debate', it is typically important to realists that this account is consilient with other theories in psychology. Psychology and psychoanalysis have the same object of study, and therefore should aim at coherence, even if they have legitimately distinct methodologies and emphases. And so we may note that this model of interpersonal mental functioning has received considerable empirical support. See, for example, the 'internal working model' of attachment theory (Cassidy and Shaver 1999), Stern's (1985) concept, in developmental psychology, of RIGs (Representations of Interactions that have been Generalized), and Andersen and colleagues' theory of the relational self in personality psychology (Andersen and Chen 2002; Andersen and Thorpe 2009). See also Narvaez, this volume.

(⁵) Mitchell and Aron (1999: xv) state that 'relational analysts, in speaking of a two-person psychology, have never intended to deny that there are two distinct individuals with their own minds, histories and inner worlds, but rather have meant simply to emphasize the emergence of what Ogden calls the intersubjective analytic third'. Mitchell (2000b: 57) expresses the emergence the other way around: 'individual minds do arise out of and through the internalisation of interpersonal fields, and … having emerged in that fashion, individual minds develop what systems theorists call emergent properties and motives of their own'. The two are compatible, as he then explains: 'Interpersonal relational processes generate intrapsychic relational processes which reshape interpersonal relational processes … '.

(⁶) Eagle, Wolitsky, and Wakefield (2001: 477) compare the debate about the priority of object relations as structures and experience to that in personality psychology between trait theorists and situationists. That debate is largely resolved in favour of interactionism, in which behaviour is a product of both trait and situation (see, e.g., Fleeson and Noftle 2009). But constructivists can complain that this analysis, in terms of how the individual interacts with the situation, is the wrong way around, at least in psychoanalysis, given that the interpersonal situation is primary.

(⁷) A connected thought here may be that psychoanalysis provides not so much a *causal explanation* of occurrent psychological phenomena as a *descriptive identification* of a pattern that may be found in them. On this view, talk of object relations is best understood not as describing 'real structures' but as an abstraction. On the importance of the distinction between description and explanation in psychiatry and psychology, see Jaspers (1913), Murphy (2007: ch. 5), Gipps (this volume), and Gipps (2018). It is notable that a considerable portion of the strength of the original argument for the hermeneutic interpretation rested on a distinction between motives and causes (e.g. Ricœur 1970: 358–63; Taylor 1964): If motives are not causes, then explanations in terms of motives, such as those given in psychoanalysis, are not causal explanations. Not only is the explanatory form different, to think of motives as causes is to reify them, and so to postulate metaphysically dubious entities. I briefly discuss the success of this argument in Lacewing (2013, 2018).

(⁸) For a dissenting view, see Gipps (this volume).

(⁹) A different reading of the debate may take the disagreement as primarily about clinical technique, dressed up as a debate about knowledge and objectivity. Given the origins of the debate as discussed in the sections 'The Analytic Relationship' and 'The Confusion over "Constructivism" ', this interpretation has considerable justification, but it requires one not to take the later arguments of the proponents at face value.

(¹⁰) Bernstein does not seek to *replace* truth as the *telos* of enquiry with something more 'practical', but argues that we must understand truth in terms of the practical character of all reason. Following Peirce, one pragmatist conception of truth is asymptotic. A claim is true if it would survive critical reflection under future improvements in our knowledge, or again if it remains a coherent part of our epistemic practices (Chang 2017; Putnam 1981; Wright 1991).

Such a conception of truth is compatible with many of the central claims of constructivism, as illustrated by the fact that it is held by the intersubjectivists Stolorow et al. (2006: 387), who say that there are 'no God's-eye views of anything or anyone. But it does not mean that we abandon the search for truth, for lived experience, for subjective reality. We hold that closer and closer approximations of such truth are gradually achieved through an analytic dialogue in which the domain of reflective self-awareness is enlarged for both participants.'

However, in light of the revised conception of rationality, we needn't suppose that we will or could ever reach such an asymptotic point—the point is 'ideal', something that regulates our language and practices. Nor need we conceive convergence of all true claims on a *single* point. We can recognize that we have multiple practices and so there is no single true description of reality (Putnam 1992).

([11]) Stern sometimes situates his constructivist claims about experience, language, and unformulated experience in psychoanalysis within the broader claim that *all* experience is constructed by language. This suffers the same drawbacks as Mitchell's reference to Bernstein's third category. If all experience is constructed by language, realism versus constructivism about the unconscious is on a par with realism versus constructivism about plants. The important point for constructivism in psychoanalysis is that language is not only the medium of knowledge (as it is universally), but constitutes the object of investigation. The role of language in interpersonally making sense of one's self in interpersonal experience is importantly constitutive in a way over and above the way it is constitutive (if it is) of all experience. Reasons of space prevent further exploration of this claim, but see Taylor (1985).

Michael Lacewing

Dr Michael Lacewing is a former Vice-Principal Academic and Reader in Philosophy at Heythrop College, London, an Honorary Reader in Research Department of Clinical, Educational and Health Psychology at University College, London, and a teacher of philosophy and theology at Christ's Hospital School, Sussex. He has published widely in philosophy of psychoanalysis, metaethics and moral psychology, alongside writing textbooks for A level philosophy and training in Philosophy for Children (P4C).

Oxford Handbooks Online

Abstract and Keywords

Conceptions of psychoanalysis as science typically construe its key formulations as providing posits to be referenced in inferences to the best explanation of the clinical phenomena. Such was Freud's own vision of the discipline he created, and it is reflected in his ambition to prove his fundamental formulations warranted. The first half of this chapter argues that this approach and this ambition significantly underestimate the significance of the psychoanalytic project. The suggestion is that it misconstrues foundational—and therefore unaccountable—forms of revelation, apprehension, poiesis, and grammar as merely factual claims, accounts, representations, and posits which, as such, now falsely appear to require scientific accounting. The second half of this chapter explores the significance of this for attempts to capture psychoanalytic theory and therapy within the 'reflective scientist practitioner' model of contemporary clinical psychology.

Keywords: psychoanalysis, psychology, revelation, representation, science, grammar, posit, poiesis

Richard G. T. Gipps

(p. 433)

> *You have a new conception and interpret it as seeing a new object. You interpret a grammatical movement made by yourself as a quasi-physical phenomenon which you are observing …*
>
> *But there is an objection to my saying that you have made a 'grammatical' movement. What you have primarily discovered is a new way of looking at things. As if you had invented a new way of painting; or, again, a new metre, or a new kind of song.*

(Wittgenstein 1953: §401)

Introduction

Dialectics versus Dogmatics

IN an unfinished work comparing and contrasting psychoanalysis with physics, Freud comments 'Psychology, too, is a natural science. What else can it be?' (Freud 1938). To which one is tempted to respond 'Well, might it not best be considered a *social* science?'

Such a temptation would be doubly misplaced. Tempting it is because there are important contexts of scientific enquiry that give sense to a distinction between the natural and social, and in such contexts psychoanalytic psychology is, despite what Freud the theoretician often suggests, far more naturally characterized as social science (see Lacewing 2013a, 2018a). Misplaced, though, because a contrast between the natural and the social is, as it happens, simply not *here* the one giving sense to his discussion. For although he soon goes on to mislead himself about what he is saying, at this point in his essay Freud's talk of a 'natural science' refers to enquiry into the *nature* or being (p. 434) of something—an investigation which is equally pursued by those social sciences whose methods have little in common with the methods of physics, chemistry, and biology.[1] And misplaced too because, in other contexts, we may wish for what I will make a case for later on: the characterization of the fundamentals of psychoanalysis as less (what could be called) science and more (what could be called) art.

I mention this here by way of introducing both my methodological and my substantive concerns. In what follows I'll be finding ways to give some determinate content to the otherwise rather vague question of how and when and whether psychoanalysis is or is not science, each time deploying a contrasting pair of terms—accounting versus acknowledging; explanation versus understanding; representation versus revelation; posits versus poiesis; cognition versus recognition—and locating fundamental aspects of psychoanalysis on the latter (let's call it 'non-scientific') side of the contrasting pair. But my purpose in doing so would be mistaken if the reader were bluntly to ask 'so are you saying that there are no respects in which psychoanalysis may be called a science?' For might we not also offer measures of scientific prowess—such as the use of clear methods, in a controlled environment, with careful observations, generating a body of knowledge, capable of inspiring testable hypotheses, consilient with extra-clinical psychological knowledge etc.—and find that psychoanalytic theory thus measured at least sometimes comes out as scientific in character? That this is so has been carefully argued by my co-editor (Lacewing 2013a). And since I'm largely persuaded I shall instead focus here only on the question of how and where psychoanalysis is less science, more art.

The substantive thrust of what I have to say is as follows. Every discipline—scientific or not—has what we might describe as its founding revelations and its revelation-presupposing representations.[2] When the revelatory character of its foundations goes unacknowledged, a standing temptation of apologists is to urge the alleged veracity and validity of what are now cast not as revelations but merely as foundational *suppositions*—while critics focus on the alleged lack of warrant for these supposed suppositions. This

has, from Freud onwards, happened frequently in relation to psychoanalysis. Yet in so doing both sides downplay the originary, *sui generis* significance and artistry of its revelations, ignoring this while recruiting extra-psychoanalytic data to frame an otiose defence or besides-the-point criticism. Regarding such manoeuvres the critic may—will, by me—be accused of having a tin ear, and the apologist of a failure of self-possession. The central tenets of psychoanalysis—that people are motivated by unconscious wishes, that dreams and parapraxes may be interpreted, that we labour under diverse defence mechanisms that preserve us from anxiety yet restrict our engagements, that our significant relations are so often replete with vexing and unwitting transferences, etc.— are (contra (p. 435) the pundits and critics) neither properly scientific nor regrettably unscientific but, I claim, non-scientific, belonging instead to the empirically unassailable artistry of its founding revelation. Indeed the claim here is that the distinctive psychoanalytic branches of representational science only have a psychoanalytic tree to hang on because of that tree's rootedness in its ownmost non-representational revelation.

There is another—methodological—reason why this chapter does not take itself to answer 'the' question of whether psychoanalysis is art or science. This is because the time is surely long since past when one can honestly say it makes good sense for philosophers to debate, say, 'whether psychoanalytic understanding treats of meanings or of causes'. This as I see it is neither because nuanced philosophical argument has finally revealed that, in fact, psychoanalytic understanding treats of one to the exclusion of the other, nor because modern philosophy of mind has cleverly shown that the two are not in truth disjunctive. It is rather because a philosopher—typically an analytic metaphysician—raising *that* question in *that* bald a manner potentially betrays a lack of dialectical sophistication which threatens to vitiate her philosophical project *ab initio*. For in a truly critical context it remains to be demonstrated, rather than simply presupposed, that there already is a determinate if yet unexplicated meaning to the terms ('causation', 'science', 'meanings', and, especially, 'the mind') of such questions in the context of their current deployment— which meaning may then be taken for granted so that something called 'working on' an 'account' that allegedly answers for us 'the problem' or 'the question' of 'the' unconscious may be undertaken.

The problem is particularly acute when it proceeds upon a foundation of careless substantialization and hypostasization—as when one prescinds from enquiry into the witting and unwitting, life-historically and socially situated, acts and habits and dispositions (call them 'comportments') of human beings, and from reflective enquiry into the ascription conditions for the various concepts for these comportments, instead focusing on what in a typical jargon may be construed as the 'states and processes' 'going on in' something called '*the* mind' or 'conscious*ness*' or '*the* unconscious'. Let's distinguish such comportments—the concepts of which may be generated by the nominalization of various verbs for human activities and actions such as 'think', 'wish', 'understand', 'feel' etc.—from those call-them-'phenomena' as are *originally* identified (referentially, for example) by nouns (like, say, 'kidney', 'corpus callosum', 'retina'). While there's surely nothing *intrinsically* wrong with nominalization, this conceptual manoeuvre by itself can hardly create a phenomenon, and a fortiori not thereby establish that the

said comportments enjoy a nature sufficient to sustain and reward the same kinds of inquiry as may be directed at genuine phenomena.

In truth, we philosophical psychologists by now know far too well where such undialectical efforts typically leave us: with increasingly complex 'explanations' of alleged phenomena that have us everywhere hanging in the air, vainly hoping our feet might one day rest back on the psychoanalytic ground if somehow we can now also find it in ourselves to commit to an ever-proliferating host of accounts issuing from sundry interdigitated inquiries in the philosophy of mind, philosophy of science, cognitive science, metaphysics, epistemology, etc. Given the difficulty of keeping so many balls in the air at (p. 436) once—without building a castle there to keep them in—the more cautious of our analytical metaphysicians will offer that they are not themselves so much making commitments as merely sketching with greater finesse where at circumscribed points such philosophical theories do and do not cohere. Yet neither commitment nor want of finesse were our original need or concern, which was rather to make more of an actual reflective acquaintance with this charming yet baffling creature—psychoanalysis— and with those forms of unconsciousness of which she provides such intriguing report. To be proffered a merely conjectural schematic of her nature is to suffer the cold shoulder.

No, what we require is not, say, a guiding philosophical presumption that terms like 'cause', 'meaning', 'art', 'science', 'real', carry a univocal and always anticipatable meaning with them into just whichever context of enquiry they happen to be deployed, so that a question like 'is psychoanalysis art or science?' may supposedly be clearly raised even if not cleanly answered. A preferable approach would be for the meanings of our core distinctions (poiesis versus posits, meanings versus causes, etc.) to be not *presupposed by*, but rather *grasped as a function of*, the discriminating work they're put to, in particular contexts of clinically grounded enquiry, as we draw distinctions to elucidate the character of psychoanalytic knowledge and its objects. That, at least, will be the angle offered here, in which explanatory theories will not be provided, for the method will instead be dialectical: distinctions will be made, antitheses will be devised not borrowed.

Acknowledging versus Accounting

A frequent Freudian refrain offers us the unconscious as a hypothesis, a hypothesis allowing us to explain phenomena which, it is claimed, otherwise must but mystify. Freud's own favourite examples are parapraxes and post-hypnotic suggestions.

Imagine: the year is 1889, the place Nancy in France, and Bernheim the physician is, assisted by Freud, conducting some experiments on hypnosis:

> . . . it is possible in the case of persons in a state of hypnosis to prove experimentally that there are such things as unconscious psychical acts and that consciousness is not an indispensable condition of activity. Anyone who has witnessed such an experiment will receive an unforgettable impression and a conviction that can never be shaken. Here is more or less what happens. The

doctor enters the hospital ward, puts his umbrella in the corner, hypnotizes one of the patients and says to him: 'I'm going out now. When I come in again, you will come to meet me with my umbrella open and hold it over my head.' The doctor and his assistants then leave the ward. As soon as they come back, the patient, who is no longer under hypnosis, carries out exactly the instructions that were given him while he was hypnotized. The doctor questions him: 'What's this you're doing? What's the meaning of all this?' The patient is clearly embarrassed. He makes some lame remark such as 'I only thought, doctor, as it's raining outside you'd open your umbrella in the room before you went out.' The explanation is obviously quite inadequate and made up on the spur of the moment to offer some sort of motive for his senseless behaviour. It is clear to us (p. 437) spectators that he is in ignorance of his real motive. We, however, know what it is, for we were present when the suggestion was made to him which he is now carrying out, while he himself knows nothing of the fact that it is at work in him.

(Freud 1938: 285)

For those unfamiliar with our unconscious life such experiments are particularly remarkable. They prove the possibility of radical dissociation between the disposition to act on and the capacity to report the fact of a suggestion. When the dissociation is such that we have action in the absence of report, we meet with what we call 'unconscious intention'. The offering of that label could be said to amount to what I am here calling an *acknowledgement* of the fact of unconscious mental life. Yet despite his frequent challenge to 'anyone in the world to give a more correct scientific account of this state of affairs [which, if they succeed, will cause Freud to] gladly renounce our hypothesis of unconscious mental processes' (Freud 1916–17: 277), nothing in what Freud reports would appear to condone a description of the unconscious intention to open the umbrella as an *explanatory hypothesis* as to why the patient opened the umbrella.[3]

Or at least … call it 'explanation' if you will. But what I'm here inviting you to do is follow me in *drawing up* a—not tracing a supposedly pre-existing—distinction between acknowledgement and explanation of the psychological facts. Where by 'explanation' I intend here precisely something which goes *beyond* acknowledgement of what happens—thereby fashioning new hypotheses regarding, for example, its extrinsic causes. For in the absence of some such distinction (between phenomenologically surveying and giving voice to the character of human behaviour, and offering hypotheses to explain this behaviour's occurrence), talk of explanation is apt to appear either vapid or hyperbolic, while in its presence such talk here is, I suggest, misplaced.

Freud did more than anyone to acquaint us with the (painful, embarrassing, regressive, violent, lustful) facts of our unconscious lives. He laid them out before us, midwife to our eager and reluctant acquaintance. At some of these we marvel, especially at first, astonished or appalled at this mysterious unavowable intentionality within even (what we had imagined to be) the more innocent dimensions of our lives. Talk of 'unconscious wishes/emotions/intentions', however, does not solve the mystery of such phenomena, but

rather dignifies them with their own place in the order of things. And as we learn to offer them such acknowledgement we also learn to stop mystifying ourselves with our continued failed attempts to assimilate them to the place occupied by other matters we already understand in other terms. We may come, too, to appreciate that our difficulty in accepting unconscious intentions was partly dependent on an unwarranted assumption: that a dissociation between what even post-dissociation we may yet be drawn to call 'action' and 'avowal' was somehow simply impossible, that this defining unity of our (conscious) intention was everywhere and always indissoluble. But, well, why should it be? The possibility may frighten us; the unity thereby disturbed (p. 438) may even run deep in our very concepts of agency and intentionality. Yet none of that makes of it a natural necessity.[4]

Reflections

Grammar versus Fact

It is to Wittgenstein that we owe both the grammatical/empirical distinction and an approach to the core propositions of psychoanalytic theory as belonging to grammar. The 'proposition "There is an unconscious" is'—writes Charles Elder (1994: 46), the author of the most sustained Wittgensteinian treatment of the grammar of psychoanalysis —'*internal to the logic* of psychoanalysis'—much as 'there is length' is internal to the logic of geometry, 'momentum is a function of mass and velocity' to physics, 'there is narrative structure' to literary theory, 'plants have leaves' to botany—rather than a psychological hypothesis to be evidenced by scientifically collated and interpreted data. 'Consider this case', says Wittgenstein:

> [We] have a general undirected feeling of fear. Later on we have an experience which makes us say, 'Now I know what I was afraid of. I was afraid of so-and-so happening'. Is it correct to describe my first feeling by an intransitive verb, or should I say that my fear had an object although I did not know that it had one? Both these forms of description can be used. (Wittgenstein 1964: 22)

To help us appreciate such a latitude's legitimacy he invites us to 'examine the following example':

> It might be found practical to call a certain state of decay in a tooth, not accompanied by what we commonly call toothache, 'unconscious toothache' and to use in such a case the expression that we have toothache, but don't know it. It is in just this sense that psychoanalysis talks of unconscious thoughts, acts of volition, etc.
>
> Now is it wrong in this sense to say that I have toothache but don't know it? There is nothing wrong about it, as it is just a new terminology and can at any time be retranslated into ordinary language…. But the new expression misleads us by

> calling up pictures and analogies which make it difficult for us to go through with our convention . . .

> (Wittgenstein 1964: 22–3)

And why should this be a problem for us? Because—and here Wittgenstein tacitly references Freud's philosophical reception:

> … by the expression 'unconscious toothache' you may either be misled into thinking that a stupendous discovery has been made, a discovery which in a sense (p. 439) altogether bewilders our understanding; or else you may be extremely puzzled by the expression (the puzzlement of philosophy) and perhaps ask such a question as 'How is unconscious toothache possible?'

> You may then be tempted to deny the possibility of unconscious toothache; but the scientist will tell you that it is a proved fact that there is such a thing, and he will say it like a man who is destroying a common prejudice. He will say: 'Surely it's quite simple; there are other things which you don't know of, and there can also be toothache which you don't know of. It is just a new discovery'.

So, Wittgenstein offers us a deflationary answer to the question 'Are we *right* to talk of "unconscious fears and desires"?' We feel an undirected sensation of fear (we might also call this 'anxiety'). If you ask us 'of what?' we may either reply i) 'of nothing', or say ii) 'I don't know'. Wittgenstein invites us to see these as equivalent: we can here say what we like, although of course it's best if we don't then go on to mislead ourselves about what we mean by what we say.

And how might we mislead ourselves about that? Well, we might think that i) and ii) contrast in that the fear in i) supposedly has no object, while ii) supposedly has an unknown object. When we consider the toothache analogy, however, we're not drawn to think there's any difference between what we're here calling 'unconscious toothache' and 'tooth decay but thankfully no toothache'. This understanding can then be usefully extended to cover our grasp of the sense of the present use of 'not knowing what we fear' and 'coming to know what we feared'. We can say either that an objectless fear gained an object, or that later experience shows that the fear was only apparently objectless.

Analogies are unhelpful when overextended. Thus a clear conceptual contrast between toothache without tooth decay and intransitive fear is that there is no transitive form of toothache; toothaches are not about anything—not even about tooth decay! Furthermore, the psychoanalytic concept of 'unconscious emotion' is most often used to denote cases in which not merely the object but also the emotion itself is unacknowledged by its subject. (Or to put it with more conceptual rigour: cases in which we are invited to yet talk of 'emotion', on the basis of the characteristic shape of the subject's expressive behaviour, despite that subject's non-acknowledgement of said emotion.) Finally, we may usefully distinguish two forms of undirected fear. In the first I later discover that I had, say, unwittingly ingested a large quantity of caffeine. In the second I later 'have an experience

which makes [me] say, "Now I know what I was afraid of. I was afraid of so-and-so happening" '. The first provides a cause but no object for the undirected fear; the latter may make us want to revise our description of it as directionless. The revision would, however, still be non-compulsory; we could still equally say that an undirected feeling of fear which yet had a certain cause but no object has *turned into* a directed feeling of fear with both object and cause.[5]

Yet, while we could say this, we don't. And this is because the grammar of the psychoanalytic unconscious has a further distinctive aspect missed by Wittgenstein's suggestion that it is 'in just the sense' in which toothache may be called 'unconscious' (p. 440) that fear may also be so called. Consider the difference between Tony's organic agnosia and Heather's psychodynamic repression.

Tony experiences a neurologically rooted dissociation between what we may well enough call his experience of fear and his ability to avow reasons for the fear. There he is, quivering yet paralysed in front of the pacing tiger, but when you ask him 'why are you quaking?' he may, despite yet being able to talk, not here have an answer. In this case we may say either that 'Tony is afraid of the tiger but doesn't know it' or that 'Tony's fears no longer have objects even though they still have causes'. Tony's unfortunate condition abrogates that unity of potential avowal and affect presupposed by our ordinary ascriptions of intentional attitudes (e.g. fear *of* tigers). But, so long as we're mindful of this, we shan't get stuck with the hope that science or philosophy will, eventually at least, tell us which ascription is here the right one.

Heather, however, experiences a defensively motivated block regarding the object of her anger. She depends on her husband and the marriage is deeply important to her but, frankly, it was obvious to everyone from day one what a bastard he could be. Obvious to all but her that is: Heather makes every excuse for his behaviour. Yet her parents start to notice: how she's getting more and more angry with her children ('she's deflecting her anger onto them' they say), how she's now broken a surprising number of his precious china figurines when cleaning ('oh, *another "accident"* ', they say), how she keeps on coming down with a headache when he wants her with him ('but I'm *not* angry with him!' she keeps telling them).

Later on Heather comes to acknowledge her anger at her husband. Such acknowledgement is what Wittgenstein tends to treat as the most salient criterion for her previously unconscious anger ('For only if he acknowledges it as such, is it the correct expression. (Psychoanalysis.)' (Wittgenstein 2005: 410)). But what this misses is the conceptual significance of the defence mechanisms—that they're not extrinsic causes, but constitutive, of the unconsciousness with which we here meet. That is to say: it was the positive presence of her defences, and not merely the absence of her acknowledgement … it was the situated inflection of her earlier disavowal, and not merely the fact of her belated avowal … which was criterial for our judgement that Heather was (not globally dyspeptic but rather) unconsciously angry with her husband.[6] These defences show themselves not simply as absences or blindspots, but rather as the positive presence of,

for example, defensiveness. Or, to move from denial to projective identification: imagine that we're called on to justify our judgement that her husband projects his guilt into Heather. What we should appeal to would not simply be his earlier perception of her as guilty plus his post-therapy tolerating and acknowledging the guilt as his own along with his ceasing to see her as blameworthy. We should, more importantly, also include his earlier tendency to act in a subtly wounded manner towards her, tacitly (p. 441) and unfairly insinuating that he has been traduced, and insinuating this in such a way as is likely to covertly induce guilty feelings within her.[7]

It's time to regroup. Wittgenstein offers us an important thought—that the existence of the dynamic unconscious belongs to the 'logic' or 'grammar' of psychoanalytic theory. The implication of this is that we don't do well to think of the existence of the dynamic unconscious as something to be justified by appeal to scientifically collected observation. It is, to use the term which in this chapter I'm borrowing from Thomas Mann's (1947) 'Freud and the Future', rather an intrinsic and unassailable part of the psychoanalytic *revelation*. Yet sometimes when Wittgenstein aims to characterize the grammatical character of core psychoanalytic tenets he does so by reducing them to phrases that may supposedly be translated out without loss into non-psychoanalytic idioms (as we might translate 'bachelor' into 'unmarried man'), rather than offering them to us as a *sui generis* revelation of aspects of human reality not otherwise visible. As he says in the *Philosophical Investigations*:

> But there is an objection to my saying that you have made a 'grammatical' movement. What you have primarily discovered is a new way of looking at things. As if you had invented a new way of painting; or, again, a new metre, or a new kind of song.
>
> (Wittgenstein 1953: §401)

Learning to hear and sing this song depends not on learning to substitute handy psychoanalytic epithets for cumbersome non-psychoanalytic phrases. It involves the development of new sensibilities, the development of a new attunement to the distinctive shape of that affective and conative domain we call the 'dynamic unconscious', an attunement required for bringing into view the facts of unconscious motivation and emotion. In justifying an interpretation a psychoanalyst may appeal to certain of her patient's expressions and behaviours, but how it is that these are to be related to their context, and what aspects of them are here the pertinent ones, is not to be understood through a reflective grasp of a general rule. It depends, rather, on the gist-getting afforded by a hard-won analytic sensibility. In this way the unconscious aspect of the situation may be appreciated in its particularity.

Revelation versus Representation

One way to draw, and thereby give sense to, a distinction between what I shall call phenomenological revelation and scientific representation is to focus on the latter's method, sometimes called *the* scientific method, of asking a question of nature in such a

way that, while *we* have set the terms of the question, *she* at least gets to answer without our further interference. So, I ask: 'Does X obtain or not?' If—and here I stipulate—my (p. 442) question is to be taken as scientific, my grasp of the meaning of 'X' must be largely independent of my understanding whether X actually here obtains. In particular it cannot be, on this stipulative understanding of a condition necessary for an enquiry to count as 'scientific', that the shape of my understanding of the meaning of 'X' simply be a function of my past and present practical engagements with Xs in the course of my life. If that were so, what role would there be for empirical enquiry? If my understanding of what it is to be an 'X' is constituted by my practical ability to engage with Xs in the course of my life, rather than by that kind of theoretical knowledge as does not bottom out thus in praxis, we should not yet have cleared the space for nature to answer in her own way those existential questions we put to her. ('I know what "humans" means, but are there any?' is nonsense; 'I know what "yeti", "photographic memory", "muon", and "human pheromone" mean but, in point of fact, do they obtain?' is not.)

I said in the Introduction that my co-editor had largely convinced me that, in the context of debates as to what kind of science psychoanalysis is, we typically do better to consider it social than natural. Here, however, is something over which we've fruitfully disagreed: 'Suppose', he writes, that:

> someone understands, theoretically, the concept of resistance. However they doubt that resistance ever occurs. What they doubt is that there *is* any such pattern in human behaviour.... Compare: a palaeontologist claims to have discovered a new dinosaur from the re-examination of existing finds. Evidence can be laid out, not by presenting new material, but by reorganizing existing material and drawing attention to what has been overlooked.... Where we argue for patterns in what we can observe, we want to know if the patterns are 'really there', in the bones or the behaviour. We are aware that we can be mistaken about this (as popular but false accounts in human behaviour and character, I suggest astrology or the four 'humours'). We need to defend the sensibility that purports to identify the pattern against the charge of confirmation bias ... This is one reason for saying that we 'posit' the pattern ...

> (Lacewing 2018b: 120–1)

Here, where Lacewing compares, I contrast. We may imagine that the posited patterns, which patterns partly constitute the meaning of the terms for the mooted species, did or did not once actually obtain. One palaeontologist may arrange the material in one way, another in another, and only one (or neither) of these scientists be positing a possibility that was also an actuality. It is natural, too, to take a similar line for the claims of astrology and humoralist medicine: their posits counting as true or false hypotheses of celestial or physiological events designed to explain the occurrence of the patient's manifest life events and symptoms.[8] Yet this, I maintain, is precisely what should be questioned in relation to the central concepts of psychoanalysis.

A New Kind of Song: Psychoanalysis as Revelation

Let's stay with the concept of resistance. What might a purely 'theoretical' understanding of it look like? Well, Freud wrote:

> The discovery of the unconscious and the introduction of it into consciousness is performed in the face of a continuous resistance on the part of the patient. The process of bringing this unconscious material to light is associated with pain, and because of this pain the patient again and again rejects it.

(Freud 1904/1959)

Neither Lacewing nor I would disagree that it's possible to learn to parrot such phrases, perhaps even write academic papers which in some sense are about psychoanalysis, without having a real grasp of their meaning. And we could even both agree that a merely 'theoretical' understanding of the term 'resistance' was possible—*if* by this was meant one that was sketchy, schematic, unfleshed out. (We may know 'theoretically' that 'resistance = voltage/current' is an equation of physics without really knowing what this equation is on about.) My claim, however, is that to actually *know what psychoanalysis means* by 'resistance', we have to learn how to use the term in practice—to become actually and imaginatively acquainted with the distinctive shape that resistance takes in our own and others' lives. (Resistance, one might say, *looks like* something, and it is this look we must learn.) And that, before we've done this, we don't really know what it means and also that, once we've done that, there's no intelligible further question for us as to whether it actually exists or not. Contrast the newly posited dinosaur which may only have been stomping about in our palaeontologist's mind and not also in the swamps of prehistory.

To rehearse, the purpose of this discussion is not to draw up an artificially hard-and-fast distinction between phenomenological elucidations and scientific posits, point out the ways in which the essentials of psychoanalytic theory are phenomenologically bent, and use this to force a conclusion that psychoanalysis at its heart is, as it turns out, not a science. Instead, I proceed in the manner I described as 'dialectical': i.e. to give content to a distinction between science and non-science by *drawing up* a living distinction, then applying this to psychoanalytical theory. The claims so far have been that 'people have an unconscious/manifest resistance/engage in projection/etc.' belongs to grammar and not science (i.e. we couldn't understand it without already grasping its truth, thereby guaranteeing its empirical unassailability and also the meaninglessness of attempting its empirical confirmation); that learning this grammar involves not learning abstract rules but becoming phenomenologically familiar with the manifestations of the unconscious; and that such learning involves the deepening of a distinctive sensibility and *sui generis* attunement to the subtle and complex, always situated and inevitably particularistic, otherwise invisible resonances of our affective and conative lives.

Now, it occurs to me that a reader not yet catching my drift may think that Mann's and my talk of 'psychoanalytic revelation' smacks of mystery-mongering. As if we were claiming that the unassailability of the core propositions of psychoanalysis were a matter of gnosis: a matter of an intuitive experiential revelation of the truth of the claims as

propounded by the doctrine. To answer I should like to emphasize a distinction between grammar and gnosis.

(p. 444) My contrast, recall, was between revelation and representation. A claim to gnosis is—let's say—a claim to an intuitive, and hence unassailable, personal access to those facts which determine the truth or falsity of our representations of them. Yet what instead I've been maintaining is that claims such as 'people sometimes repress their emotions' don't articulate what I'm here calling representational knowledge at all. For whereas the stipulated hallmark of such knowledge is a disjunction of matters of meaning and truth, our coming to understand the truth of such fundamental psychoanalytic propositions is of a piece with our coming to understand their meaning. Or to put it otherwise: to say 'people sometimes repress their emotions' is of the order of 'sometimes it rains', or 'human life is permeated by power relations', or 'people typically prefer requited to unrequited love', or 'we tend to try to get or achieve what we desire'. On the one hand: empirical facts. On the other: not something anyone understanding how to use these words aright could meaningfully doubt. If you want to learn what is meant by 'weather': come outside! By 'power relations': get yourself sensitized through a sociology class and micropolitical discussion. By 'unconscious affectivity': get analysed, search your heart, keep a journal, do an analytic training, read literature. In all such cases our theory is necessarily data-laden from the get-go, yet the data are themselves what are disclosed by the same world-opening revelation as is articulated in the theory.

In truth it is not the revelatory, but the representational, vision which here makes for the greater mysteries. For, according to the view of psychoanalytic concepts as referencing posits intelligible independently of that relational behaviour which they are to explain, we are to grasp their meaning—not through gradually getting our psychodynamic eyes in, in the midst of our relational lives, but rather—by pulling it out of our disengaged noggins in mysterious acts of pure intellectual intuition. The motivation for the representational view may be the wish to scientifically demonstrate and justify the truth of psychoanalytic claims to one's own and doubters' satisfaction. But what we lose by abandoning it, by way of a hope of provability, we gain by way of an appreciation of the psychoanalytic revelation's originary significance.

Poiesis versus Posit

The philosopher best known for his determination to recover the disclosive revelations latently underpinning the representational edifices built on top of and, as the metaphors die, disguising these foundations is Heidegger. To briefly rehearse his idiom: 'adequatio'—representational truth—presupposes 'alethia'—a primordial uncovering or expressive disclosure of aspects of reality within the 'lichtung' ('clearing') of our worldly engagements (Heidegger 1927/1962: section 44). Such primordial disclosure is, in the philosophical context, the task of an 'inceptual' or founding phenomenology, and the 'logos' of this phenomenology is not descriptive accuracy but evocative potency or 'poiesis' (Heidegger 1989/2012: sections 20–31). As the name suggests, poetry provides its best model, since it alone finds the words with which phenomena may sing to us with

their ownmost voice, rather than that voice getting drowned out by our attempts to sing

(p. 445) of them in a voice not their own. Poiesis, on this understanding, is no mere colourful gloss overlaying the factive but rather its condition of possibility, and if matters appear otherwise that is a function of the deadening of the metaphors constituting the poietic disclosure.

I mention this by way of introduction not to Heidegger but to the philosophical significance of psychoanalytic style. Much has been written by psychoanalysts about their practice as an art which nourishes a process of self-becoming in the subject itself best articulated in aesthetic terms (Hobson 1989; Ogden 2005; Summers 2013), and much too by literary critics on psychoanalytic theory as itself an art (Harris Williams 2010; Jacobus 2005; Meisel 1981). What I wish to draw out here are those aspects of this aesthetic valuation which speak to its ontological revelation of unconscious life.

Many have acknowledged the power of Freud's (e.g. Gay 1978) and Winnicott's (e.g. Ogden 2001) style, and it is easy to contrast it with, say, the written inelegance of Pierre Janet, Carl Jung, and many an ego psychologist. Those who have read such post-Kleinians as have deepened our access to the psychotic mind (Bion, especially) will also have been struck by their confounding style's disturbing evocation of mental pain (Ogden 2005; Rosegrant 2012). The angle being developed here offers the significance of such 'style' as best grasped not by a dialectical contrast with 'substance', of which it is (in the sense offered here) in any case not the garb but the formal condition, but by constitutive contrast with the efforts of one of human nature's less successful suitors. Freud's status as a central pioneer of the territory of the dynamic unconscious has itself, on this understanding, to do with his capacity to offer it linguistic emancipation and voice, something for which flatfooted scientific forerunners of the 'unconscious idea' had little aptitude, by contrast with its prefigurement by the expressive writings of the poets and philosophers to whom Freud himself offered acknowledgement. This thought, I should acknowledge, has already been developed by Lacan (1977: 144): 'Freud ... derived his inspiration, his way of thinking, and his technical weapons' from a study of the resonances and significations of art, and also by literary critic Lionel Trilling (1950: 52–3): 'the Freudian psychology ... makes poetry indigenous to the very constitution of the mind ... [by contrast with the view] that poetry is a kind of beneficent aberration of the mind's right course'. The thought is not one simply of, but also one reflexively about, psychoanalysis, for 'Freud has not merely naturalized poetry; he has discovered its status as a pioneer settler, and he sees it as a method of thought'.

Yet for all this talk of psychoanalytic poiesis, what surely often takes us aback is how unpoetic it is with its proliferation of organismic and mechanistic tropes. To be sure we do meet from time to time with, say, 'holding', with 'dreaming undreamt dreams', with 'unthought knowns'. But the mainstay metaphors are of 'objects' (internal, transitional, part, whole, good, bad) and their 'uses', and of the properties of an internal 'container' called a 'psyche/mind' with (inter alia) its ego capacity, libidinal energy, links, libidinal cathexes, contact barrier, ego, superego, id, preconscious, conscious, unconscious, implosion, petrification, engulfment, projections, and transferences. Perhaps this should

be no surprise when we contemplate how the conceptual metaphors for our mindedness which psychoanalysis extends are themselves largely based on the body—attitudes, (p. 446) under-standing, attraction and repulsion, feeling up or down, high or low, crushed or shattered or going to pieces—or bodily contents—with feelings that bubble or are bottled up, pride swallowed, emotions spilling out, being under pressure, steam let off, cooling off, blood running cold, blowing our top (Hopkins 2000). Perhaps we should acknowledge, too, that the blatantly physicalistic bent of psychoanalytic discourse too often feeds a sterile, arcane, and pseudoexplanatory 'metapsychology' when the metaphoric music dies. Yet I here wish to make the plea that its constructive ends be kept in view too. Methodologically, what to anyone not lost in metapsychological self-satisfaction appears its flagrantly metaphoric character keeps inceptual thought alive, since the terms' blank corporeality forces their deployer to continually call up his analytic sensibility to animate them afresh. Experientially, it keeps us close to the life of the individual human being. And ontologically it keeps close the character of unconscious life, with its as-yet unsymbolized, unowned, and unsubjective bent. Psychoanalysis, after all, is not itself primarily tasked with the voicing of our better selves, but with rescuing for us our precipitately disowned feet of clay.

Having now pulled into view something of what is meant by the psychoanalytic revelation, I turn to some applications, asking of the significance of this reflective understanding of psychoanalysis for psychological practice.

Applications

Understanding versus Explaining

Philosophical psychology was once conducted through elaborating contrasts—between reasons (or meanings) and causes, between explanation and understanding. This provided a nice opportunity, with varyingly successful results, for later philosophers to try their deconstructive hands—by showing for example that some reasons are, *after all*, a species of cause. I mention this here by way of reminder that, while I shall now talk of explanation versus understanding, I will have failed myself if I trade on prior and independent understandings of explanation and of understanding. Instead I shall proceed in that manner I have described as 'dialectical'—i.e. by *not tracing but illustratively drawing* a living distinction and using 'explanation' and 'understanding' to index it.

So, consider the kinds of psychoanalytic claims which self-styled scientific psychologists tend to find perversely unevidenced. The psychoanalyst David Bell presents us with the psychoanalytic understanding of paranoia, and tells of how a:

> … threatening mental content is transported from its internal position and relocated in the external world … [N]ow objects in the external world appear to the individual to be increasingly menacing … It is this psychological mechanism that underlies all paranoid mental states.

A psychiatrist on an acute admission ward brushed past a male patient. The patient turned to him, his face a mixture of anger and anxiety and exclaimed, 'I am not (p. 447) homosexual.' How is one to understand this event? … The man … might have perceived within himself a homosexual impulse that was pressing forward for awareness. But, for whatever reason, this was intolerable and so could not be admitted into consciousness. He defended himself through projecting the impulse, and the awareness of it, into … the psychiatrist. Instead of feeling that a disturbing idea was forcing itself upon his own mind, he believed that the idea belonged to an external figure, the psychiatrist, who was trying to force it upon him.

(Bell 2003: 3–5)

Throughout his book Bell provides further Freudian theory concerning the relation between paranoia and sexual and aggressive inner conflict; details the relief to the patient resulting both from their projective disowning ('these are *his* thoughts not mine!') and also from the illusory clarity provided by paranoid delusions ('*now* I know what's going on!'); and further details the Kleinian understanding of the relation between the intolerance of ambivalence, splitting and projecting, paranoid fear, reintrojection, further projection etc. What I should like to stress here is that Bell provides evidence for these claims only in the sense in which a case study provides evidence for a theory: the study helps us understand *what* is meant by the theory by *exemplifying and enlivening* its terms; yet in its singular nature it naturally does nothing to let us know how often what is being articulated obtains. We could call this method 'phenomenological' in the way it unpacks for us a form of understanding, using the singular to give sense to—but not to evidence—the general.

Contrast the cognitive model of paranoia offered by the psychological theorist Richard Bentall which has it that:

. . . the paranoid world view arises from the tendency to make extreme self-protective attributions, together with the failure (for whatever reason) to take into account situational causes of events.

(Bentall 2003: 343)

By contrast with Bell's phenomenological approach, Bentall's empirical method aims not to elucidate for us forms of understanding, but to scientifically demonstrate that what we already understand is in fact instantiated in the mind of the paranoiac. Thus his theory identifies the possible *external triggers, internal states*, and *internal traits* which may conceivably give rise to the paranoia (the explanandum). The latter two—the inner states and traits—are that which our psychological explanation appeals to (the explanantia). Measures are taken of paranoia and of these inner states (degree of implicit low self-esteem—how the person deep down feels about or represents herself to herself; degree of explicit self-esteem—how the paranoid person consciously and explicitly represents himself to himself), and inner traits (habits of information processing such as having a

bias towards making external and personal attributions over the causes of the triggers). The measures are then correlated. A positive relation between the measures of the explanantia (the degree of implicit low self-esteem, the attribution bias) and the measures of the explanandum (the paranoia) is found, and this is taken to constitute evidence for the truth of the psychological model. The character of the theory might be summed up like this: paranoid people are also *people like this*: ... perhaps it is in part (p. 448) *because they are like this* that they are now *paranoid*; the data collected provide *empirical evidence for the truth of* the theory.

Now, Bentall's method runs into various self-confessed difficulties around testability (Bentall 2003: 339)—perhaps because it is (I suggest) impossible to meaningfully operationalize, or because it is (he suggests) hard to accurately test for, underlying as opposed to explicitly expressed low self-esteem. It also leaves an impossible-to-close gap between matters correlational and causal: all the evidence in the world that people with attribution bias and low self-esteem tend to form paranoid delusions could not establish that the latter is an effect of the former. But what I should like to focus on here is what may rather appear to be the superiority of what I'm calling Bentall's 'empirical' method over Bell's 'phenomenological' approach. For whereas Bell makes claims about paranoia but offers no evidence of their actual truth—beyond a few case studies which for all we know may not be representative—Bentall takes the time to take many careful independent measures of paranoia and of the low self-esteem and externalizing bias which, he proposes, cause it. Does this not make for the greater respectability of Bentall's method?

If it were the case that psychoanalytic claims about paranoia—concerning the anxiolytic reshaping of the person–world encounter by the projection of unthinkable internal horror into the thinkable world beyond—were intelligible independently of the kind of evidence which might be provided for them, then we can readily see how an empiricist approach would here have the upper hand. For then we should go about devising separate measures of projection, of thinkability, of paranoia, and trouble ourselves to assess their correlations. We could, as I have been putting it, offer a determinate question to nature and then await her answer. Yet we have no reason to think that these phenomena are separable, and so no way here of questioning nature, as opposed to simply listening to the sound of her voice. In fact the understanding I here offer as to what the psychoanalytic theory proposes is precisely that the phenomena are not separable: the claim being not that the distress of paranoid people is as some mere matter of fact typically metabolized by projection, but instead that projection here *characterizes* the paranoid reaction (i.e. it is its formal, rather than efficient, cause). Or, as we also might gesture towards the conceptual distinction at issue, here we do not have to do with an *explanation* of one phenomenon (paranoia) as precipitated by some quite other phenomenon (projection), but rather with the possibility of *understanding* paranoia as itself an instance of projection-in-action.[9] The understanding tracks paranoia's motivations rather than its precursor precipitants—its pulls rather than its pushes, if you like—and thereby gives us

the means to see more deeply into it. To see more deeply not into how else paranoid people happen to be, but into what it means to be paranoid.

(p. 449) Our central thesis, to recall, was that psychoanalysis' pronouncement regarding the desire-instantiating character of psychiatric symptomatology arises from phenomenological rather than any empirical investigation and is expressed in the revelatory rather than the representational moment of psychoanalytic theory. To say that human behaviour is motivated is, then, not to state a contingent fact about it but to grasp it for what it is. You withdraw your hand from a hot stove. I don't understand this by separately identifying your action, your pain, and your desire to relieve pain, then bring them together with a 'because' to give 'she withdrew her hand because the stove was burning it painfully'. We may, of course, imagine strange cases (someone wants to burn his hand to win a perverse dare, but reflexively withdraws it from the hotplate to shoo a vexatious wasp) but the fact that we must trouble to imagine special cases itself reinforces our appreciation of the fact of the default presumption that a hand withdrawn from the stove is a hand withdrawn in avoidance of pain and burning. Motivation is what we may call the intelligible form of human action as such, and not merely its contingently obtaining precipitant.

Psychoanalysis extends this vision of orectically intelligible form into the psychiatric domain of what otherwise show up merely as symptoms. To help the novitiate horse see what it is getting at, it gives it evocative case studies, and thereby leads it to that water which yet no amount of empirical data can force it to drink. That this is so we can see even in Bentall's psychological researches: we could imagine an empirical demonstration of a perfect correlation of paranoia with low implicit self-esteem and an externalizing personalizing attribution bias. Yet to know that these are in fact its causes we require something more—not an apprehension of some further correlated fact providing evidence that paranoia is an *effect* of low self-esteem, but rather a formal grasp of what it is for paranoia to be an *expression* of low self-esteem. What this grasp amounts to may be discerned from the fact that no further empirical datum—nothing from the brain and behavioural sciences for example—could conceivably provide it, for any such further psychological or neurological discovery would simply hang alongside what we already have as yet another correlated finding. Instead the phenomena of paranoia and the mental pain of low self-esteem are brought together under the motivational concept of projection. It is motivation's 'in order to' which structures even Bentall's account, for that account's intelligibility itself relies on our grasping that the patient, say, attributes externally rather than internally because of being motivated so to do—thus: 'projection'. If we understand explanation as a requirement of scientificity then we may say that Bentall's and not Bell's account is the scientific one, yet from this we may conclude not that Bell's account is unscientific but only that, in the particular sense delineated here, it is non-scientific.[10]

In a second application I turn now from psychopathological to psychotherapeutic theory.

(p. 450) Recognition versus Cognition

Psychoanalytic psychotherapists are sometimes criticized for offering patients nothing but new *Just So* stories in the guise of applied psychological science. These supposedly explain the origins of troubles in a manner which is either relieving because spuriously absolving ('it wasn't *you*, it was your *unconscious*/your *mum and dad*/your *past traumas* …' etc.) or because spuriously hope-engendering (the irrational hope being that mere reflection on your unconscious motivation will somehow help it change). The criticism continues that such a practice is: deluded since the causal stories we learn to tell about our symptoms are nothing but post-hoc fabrications; dependency-promoting; and largely ineffectual since it is concerned with introspection rather than change (Watters and Ofshe 1999). Psychotherapists' responses to such critiques vary from bite-the-bullet it's-all-just-a-story-anyway postmodernism, to that of the reflective scientist practitioner who draws as far as possible on objective psychological knowledge while modestly refraining from offering anything other than flexible revisable hypotheses in a pragmatic fashion in his clinic.

A shared assumption of both the critic and the pundit is that the psychotherapeutic work of 'making the unconscious conscious' essentially involves helping a patient arrive at new beliefs about the history and current operations of his psyche to further his self-understanding and self-management. In what follows I suggest that this 'applied science' conception locates the therapeutic endeavour in largely the wrong conceptual context. In short it locates it within what we could call an empirical psychology that offers *cognition*, rather than a moral psychology that offers *recognition*. (In line with this chapter's tack the claim is that the latter and not the former context is the more fundamental one, since it alone enables the encounter within which minded human beings show up as such.) What follows provides the substance to my contrast. I start with considering the therapist's role in making her patient's unconscious conscious before considering the patient's task of knowing his own mind.

By way of an example of an empirical psychological treatment offering what I'm calling cognition consider the following from a proponent and a critic of what both authors call 'psychodynamic psychotherapy'. First the proponent (Cabaniss et al.):

> A psychodynamic formulation … is an hypothesis about the way a person thinks, feels, and behaves, which considers the impact and development of … thoughts and feelings that are out of awareness—that is, that are unconscious…. Thus, a psychodynamic formulation is an hypothesis about the way a person's unconscious thoughts and feelings may be causing the difficulties that have led him/her to treatment…. [H]elping people to become aware of their unconscious thoughts and feelings is an important psychodynamic technique…. Once we have a good sense of the problems and patterns, the next step in creating a psychodynamic formulation is to review the developmental history…. Having described and reviewed the patient's problems and history the third step is to 'link' them together. [This provides the psychological 'hypotheses' which help the therapist

to] construct meaningful (p. 451) interventions.… These might include: … creating a life narrative … offering explanation and perspective throughout the therapy … consolidating insights …

(Cabaniss et al. 2013: 3–238)

Now the critic (Watters and Ofshe):

> Psychodynamic therapists claim the ability to help clients connect current behaviors to long-past traumas in childhood, for instance, or to repressed fantasies decades in the patients' past.… But … if [as we have shown] we can't trace the influence of simple actions and decisions to their correct sources, can we be expected to do better making etiological connections between complex current life and events or fantasies from our childhood? … [T]he vast number of psychodynamic schools of talk therapy appears as nothing more than a testing and breeding ground for these shared cultural narratives. Psychodynamic therapy offers a new and interesting world of possible narratives by which patients can come to *believe* they understand the origin of their thoughts and behaviors. These narratives become plausible in the patient's eyes through the process of influence embedded in therapy.
>
> (Watters and Ofshe 1999: 204)

Now, in what follows I don't so much argue against a version of psychotherapy which formulates then offers interventions, explanations, and advice to the patient, as aim to recover, in line with this chapter's ambition, the therapeutic significance of the psychoanalytic revelation itself before any such representational knowledge comes on the therapeutic scene.[11]

Consider: Geraldine comes to therapy with mental and physical suffering she does not want. She may or may not have a theory as to its causes. What however she cannot see is how it already embodies the intelligible form of an emotional response to her situation. Perhaps she hopes that her therapist will somehow help her remove what she herself experiences merely as symptoms. Or, failing that, perhaps the therapist can at least offer her an explanation as to where these symptoms come from and what maintains them so she can learn to manage them better. This gives us one possible meaning of what it could be to engage in the therapeutic task of 'making the unconscious conscious': i.e. to become more *conscious of* what has been going on in the unconscious mind.

Luckily for Geraldine she encounters a psychoanalytic psychotherapist who can offer a different, deeper kind of relief. For rather than set about trying to understand and share his understanding of her unconscious *mind*, he instead offers *her* understanding. That is to say, he doesn't so much offer her causal explanations of her symptoms, but instead refuses to see them as symptoms in the first place. His analytic eyes recover what she experiences as symptoms as instead living (if undeveloped) moments of her (p. 452) subjectivity—as part of a humanly intelligible response to a world. Where she experienced

baffling anxiety, he sees her ambivalent love and hate; where she suffered from fatigue and demotivation, he sees her actively shutting herself down to avoid distressing confrontation. In his eyes his patient becomes an agent, and her symptoms dissolve into actions, reactions, and ordinary emotions. The result of her internalizing such recognition is not that she now understands her present symptoms better, but that such symptoms themselves transmute into live intelligible emotional registrations of her situation. Geraldine may have come in looking for a therapist's cognitive insight into her own symptoms, but she comes out having experienced his recognition—not of her symptoms but of her in her suffering.

It's worth noting that during the course of therapy Geraldine finds that receiving her therapist's recognition is not always a comfortable matter. For although in his acceptance of her he shows a kindness she has hitherto been unable to offer herself, his seeing her symptoms as instead actions and expressions also uncomfortably highlights her own disowned responsibility. In short, there are times when, rather than being the victim of a depressive symptom, and so unable to rise to the responsibilities of everyday life, the symptom itself is seen as her culpable refusal to engage with her predicaments. The therapist offers her conjoint acceptance and accountability. This is his 'tough love', and this is his contribution to 'making the unconscious conscious'. Through this process Geraldine doesn't become conscious *of* her previous or current unconscious mental states. Instead *they themselves* become conscious as they are restored and owned as living moments of her subjectivity, will, and accountability.[12] The critic who, at the beginning of this section, suspected psychoanalytical therapy of creating exculpatory psychologizing stories for what is in fact culpably bad behaviour would, in this situation, seem to have matters quite back to front.

In this way the correlate of her therapist's humane recognition is Geraldine's appropriation of her symptoms. She is no longer possessed by them, for instead she is now self-possessed. She 'becomes herself', as we say: she now inhabits ('integrates') and develops that which previously in her was but a latent possibility. She comes to 'know her own mind', not in the sense of knowing facts about what she thinks—for that would still leave in her a merely external 'managing' relation to herself—but in the sense of consciously developing clear opinions which guide her thought and action. Her feelings become conscious—their form (Lear, this volume) changes so that they now manifest an avowable (Finkelstein, this volume), integrated (Leite, this volume) perspective, and she may now also question the understanding of the meaning of her situations which they provide, take responsibility for them, appreciate their partiality etc. Far from possessing more psychological knowledge about, or to be in some greater kind of (p. 453) comprehending relation to, her own psychological functioning, she in fact needs rather less of these. For now, when we encounter her, what we meet with is rather less by way of her *psychology* and more by way of her *personhood*. To borrow Buber's well-known terminology, neither Geraldine nor we need now engage with her mind in I–It relations. Instead we relate to her as to a You, and it is as a more fully integrated You that she responds to us.

It is not uncommon to hear psychodynamic transference interpretations described as 'hypotheses' (see the earlier quote from Cabaniss et al. 2013). The concept alludes both to an empirical scientific attitude and also to the difference between the dogmatic and the respectful therapeutic word. In this section I've argued that what the patient needs is not primarily empirical hypotheses about his symptoms but rather his therapist's ethical recognition. To conclude it I should like to suggest that the notion of a tentative hypothesis also fails to aptly characterize the difference between dogmatic and respectful therapeutic interaction. For while it's true that the therapist may not impose her will if she's to do anything that could even *count as* offering the patient recognition, we can also readily think of occasions in which she must, in order to be truly respectful, take a strong stand against her patient's defences and be firm and frank about the conscious and unconscious intention, affect, and defence clearly manifest by him. Occasions, that is, in which to do anything other than offer the *mot juste* in forceful terms—to honestly call the patient on his irresponsibility within the therapeutic relationship, for example—would be to fail to respect him as a human subject. What this call requires by way of tempering is not hypothetical tentativeness but instead, when sincerely requested, true forgiveness.

Conclusions

This chapter began with a distinction between dialectical and dogmatic thought. Dialectical thought, as I have developed the idea, is thought which does not trace pre-existing distinctions, so does not make assailable truth claims about what independently of it is the case, but rather draws distinctions afresh within its contrastive terms of art. Its artistry may be impugned or vaunted; its veridicality may not—since it neither enjoys nor wants for it. Distinctions made along the way, within a master distinction of revelation and representation, have been that between: grammar and fact, poiesis and posit, disclosure and description, acknowledgement and accounting, understanding and explanation, phenomenology and empirical science, and recognition and cognition. In each case the claim is that, as the terms are here intended, psychoanalysis suffers a failure of nerve, and its critics a tin ear, to the extent that they take the latter, rather than the former, to provide the model for the central tenets of its theory. The intent of my enquiry has not been to suggest that the representations—i.e. the facts, posits, descriptions, accounts, explanations, hypotheses—do not have their place within what may comfortably be called the social science of psychoanalysis. It has instead been (p. 454) to caution against the fantasy that such valuable representational attic rooms could ever usurp the role of the revelatory foundations of the psychoanalytic edifice.

To close I should like to comment on the relationship between this dialectical approach to the philosophy of psychoanalysis and the ethic of psychoanalysis itself. A key theme of contemporary clinical psychoanalysis has been the significance of *keeping inner experience and thought alive*. Think, for example, of Bion on the value of 'suffering' rather than defensively 'constructing' our experience, Winnicott on the emotional significance of 'play', Bollas on the importance of 'cracking up' in unconscious experience, Ogden on 'dreaming undreamt dreams and interrupted cries', Meltzer on the

developmental significance of 'beauty', Stern on the therapeutic significance of 'vitality affects'. Such (romantically inspired) clinical approaches take an increasingly particularistic approach to psychoanalytic theorizing: we are encouraged not simply to understand the particular by bringing to bear upon it a general metapsychological concept, but to work at reversing the trajectory and keep the meaning of our theoretical terms alive by anchoring them in evocatively articulated and empathically grasped case material. Accountability, in the most fundamental junctures of both clinical and theoretical contexts, amounts here not to a demonstrable fidelity to rules or evidence; that would be the derivative, representational accountability met with in the less foundational tiers of psychoanalytic science. Instead it amounts to taking responsibility for undertaking one's own authentic poietic disclosure of the inner life. The 'new kind of song' which is psychoanalysis is not one to be learned and then formulaically repeated, but rather one to be sung afresh in the diverse changing contexts of our developing lives. Clinically this amounts to coming alive again as a human subject; philosophically it amounts to the phenomenon we call 'having something to say'.

References

Bell, D. (2003). *Paranoia*. Cambridge: Icon Books.

Bentall, R. (2003). *Madness Explained: Psychosis and Human Nature*. London: Penguin Press.

Cabaniss, D., Cherry, S., Douglas, C., Graver, R., and Schwartz, A. (2013). *Psychodynamic Formulation*. Chichester: Wiley-Blackwell.

Elder, C. (1994). *The Grammar of the Unconscious*. University Park, PA: Pennsylvania State University Press.

Freud, S. (1904/1959). 'Freud's psycho-analytic procedure". In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 7 (pp. 249–270). London: Hogarth Press.

Freud, S. (1913/1955). 'The claims of psycho-analysis to scientific interest'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 13 (pp. 163–190). London: Hogarth Press.

Freud, S. (1916–17). *Introductory Lectures on Psycho-Analysis*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vols 15 and 16. London: Hogarth Press.

Freud, S. (1938). 'Some elementary lessons in psycho-analysis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 23 (pp. 281–286). London: Hogarth Press. (p. 455)

Gay, P. (1978). *Freud, Jews and Other Germans: Master and Victims in Modernist Culture*. New York: Oxford University Press.

Hacker, P. (2002). 'Is there anything it is like to be a bat?' *Philosophy 77*: 157–174.

Hai, H. (2014). *The Theory of Chinese Medicine: A Modern Interpretation*. London: Imperial College Press.

Harris Williams, M. (2010). *The Aesthetic Development: The Poetic Spirit of Psychoanalysis*. London: Karnac.

Heidegger, M. (1927/1962). *Being and Time*, trans. J. Macquarrie and E. Robinson. Oxford: Basil Blackwell.

Heidegger, M. (1989/2012). *Contributions to Philosophy (Of the Event)*. Bloomington, IN: Indiana University Press.

Hobson, R. (1989). *Forms of Feeling: The Heart of Psychotherapy*. London: Tavistock/ Routledge.

Hopkins, J. (2000). 'Psychoanalysis, metaphor, and the concept of mind'. In M. Levine (ed.), *The Analytic Freud* (pp. 11–35). London: Routledge.

Jacobus, M. (2005). *The Poetics of Psychoanalysis: In the Wake of Klein*. Oxford: Oxford University Press.

Lacan, J. (1977). *Écrits: A Selection*. New York: Norton.

Lacewing, M. (2013a). 'Could psychoanalysis be a science?' In K. W. M. Fulford, M. Davies, R. G. T. Gipps, G. Graham, J. Sadler, G. Stanghellini, and T. Thornton (eds), *The Oxford Handbook of Philosophy and Psychiatry* (pp. 1103–1127). Oxford: Oxford University Press.

Lacewing, M. (2013b). 'The problem of suggestion in psychoanalysis: An analysis and solution'. *Philosophical Psychology 26*: 718–743.

Lacewing, M. (2018a). 'The science of psychoanalysis' *Philosophy, Psychiatry, and Psychology 25*(2): 95–111.

Lacewing, M. (2018b). 'Evidence, inference and causal explanation: a reply to Gipps'. *Philosophy, Psychiatry, and Psychology 25*(2): 119–122.

Mann, T. (1947). 'Freud and the future'. In *Essays of Three Decades*, trans. H. T. Lowe-Porter (pp. 411–428). New York: Alfred A Knopf.

Meisel, P. (ed.) (1981). *Freud: A Collection of Critical Essays*. Englewood Cliffs, NJ: Prentice Hall.

Ogden, T. (2001). 'Reading Winnicott'. *Psychoanalytic Quarterly 60*(2): 299–323.

Ogden, T. (2005). *This Art of Psychoanalysis: Dreaming Undreamt Dreams and Interrupted Cries*. Hove: Routledge.

Rosegrant, J. (2012). 'Why Bion? Why Jung? For that matter, why Freud?' *Journal of the American Psychoanalytic Association 60*(4): 721–745.

Summers, F. (2013). *The Psychoanalytic Vision: The Experiencing Subject, Transcendence, and the Therapeutic Process*. London: Routledge.

Trilling, L. (1950). 'Freud and literature'. In *The Liberal Imagination* (pp. 34–57). New York: Viking Press.

Watters, E. and Ofshe, R. (1999). *Therapy's Delusions: The Myth of the Unconscious and the Exploitation of Today's Worried Well*. New York: Scribner.

Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Basil Blackwell.

Wittgenstein, L. (1964). *The Blue and Brown Books*. Oxford: Basil Blackwell.

Wittgenstein, L. (2005). *Big Typescript (TS 213)*, trans. C. G. Luckhardt and M. E. Aue. Oxford: Blackwell Publishing. (p. 456)

Notes:

([1]) It is worth noting that Freud (e.g. Freud 1913/1955) also includes *philosophy* among 'the sciences'. While *Wissenschaften* is the general term for knowledge construction (an archaic sense of 'science'), and while Freud does sometimes use *Naturwissenschaften* which implies a contrast with *Geisteswissenschaften*, at this point in this essay in English his talk of a 'natural science' does not invoke a contrast with the social sciences.

([2]) I give specific content to 'revelation' and 'representation', and distinguish my use of these terms from other possible uses, in the section on 'Revelation versus Representation'.

([3]) This theme is developed in detail by Elder (1994: section 1.17 et passim).

([4]) As opposed to a conceptual necessity in the normal (i.e. 'conscious') cases of intentional action; see the following section on 'Grammar versus Fact'.

([5]) See the final section of Lacewing (this volume) for discussion of indeterminacy and the unconscious.

([6]) Freud (1916–17), by contrast, provides a range of criteria for an unconscious motive—from the sincere subject's acknowledgement, to his suspiciously energetic denials, to the entire psychological and interpersonal situation of the symptom or slip.

([7]) For a contrastive account which does not bring the positive presence of defences into our understanding of the very unconsciousness of such emotions and motivations see Finkelstein (this volume).

(^{8}) Yet other understandings of humoral and related systems are available—for example Hong Hai's (2014) construal of 'qi' and 'phlegm' and 'vital organs' not as hypothesized underlying entities but as heuristics and constructs to guide therapy.

(^{9}) Bentall's difficulties in empirically confirming his thesis of a connection between implicit low self-esteem and paranoia can, I propose, be understood along similar lines. Once related through the concept of projection, paranoia becomes intelligible as one of the forms that low self-esteem takes, and the specification that this low-self esteem is unconscious or 'implicit' is a way of articulating the nature of this form. The recognition of this form is not however something which can be secured through the provision of any number of further, isolable, data.

(^{10}) In his chapter Eagle (this volume) bemoans the untestability of much post-Freudian theory. My suggestion is that this may be because it's concerned more with articulating intelligible forms than with offering hypotheses about efficient causes.

(^{11}) For an alternative rebuttal of this version of the psychodynamic model which distinguishes between what is helpful in clinical theory and what is unhelpful by way of metapsychological and aetiological speculation see Lacewing (2013b).

(^{12}) Finkelstein (this volume) outlines in his own terms the contrast between what may be described as transitive (consciousness *of* something) and intransitive (conscious simpliciter) senses of 'consciousness'. To describe someone as conscious (awake) or unconscious (e.g. asleep), or a mental state as conscious or unconscious, is to talk of forms of 'intransitive consciousness'; see Hacker (2002). Lear (this volume) outlines what I here offer as a broadly ethical reading of what it is for id to become ego or for the unconscious to be made conscious.

Richard G.T. Gipps

Dr Richard Gipps is a clinical psychologist in private psychotherapy practice in Oxford, UK, and an associate of the Philosophy Faculty at the University of Oxford. He convenes the Philosophy Special Interest Group of the Institute of Psychoanalysis, the Oxford Interdisciplinary Seminars in Psychoanalysis, and the Making the Unconscious Conscious seminar series. His research interests lie in psychoanalysis, psychosis, existential phenomenology, and Wittgenstein.

Oxford Handbooks Online

Body Memory and the Unconscious 🔓

Thomas Fuchs

The Oxford Handbook of Philosophy and Psychoanalysis
Edited by Richard G.T. Gipps and Michael Lacewing

Abstract and Keywords

In traditional psychoanalysis the unconscious was conceived as a separate intra-psychic reality, hidden 'below consciousness' and only accessible to a 'depth psychology' based on metapsychological premises and concepts. In contrast to this vertical conception, this chapter presents a phenomenological approach to the unconscious as a horizontal dimension of the lived body, lived space, and intercorporeality. This approach is based (a) on a phenomenology of body memory, defined as the totality of implicit dispositions of perception and behaviour mediated by the body and sedimented in the course of earlier experiences. It is also based on (b) a phenomenology of the life space as a spatial mode of existence which is centred in the lived body and in which unconscious conflicts are played out as field forces.

Keywords: unconscious, phenomenology, body memory, life space, field forces, intercorporeality

> *... il n'y a pas d'homme intérieur, l'homme est au monde, c'est dans le monde qu'il se connaît.*
>
> Maurice Merleau-Ponty

Introduction: Psychoanalysis and Phenomenology

Psychoanalysis and phenomenology, two theories that arose at more or less the same time,[1] both considering themselves basic sciences of subjectivity, have nevertheless remained foreign to one another. The grounds for this are probably to be found primarily in their conflicting views of the role played by consciousness. To psychoanalysis, consciousness appeared only as a shimmering varnish concealing psychological forces and processes in unfathomable depths which are what is actually effective. For phenomenology, on the other hand, consciousness rather was the medium or the light through which all phenomena come to be seen in the first place, and appear as such. Consciousness as the sphere of mere *semblance* (*Schein*) or of *manifestation* (*Erscheinung*)—is a pointed distinction that could be made between the two. Accordingly, they held contrasting views also of the unconscious: either it was considered the actual source of the psyche's life, the hidden meaningful structure and driving force, which made its way by various but coded means, even in opposition to the conscious intentions of the subject. Or the unconscious had to be viewed as restricted to an implicit awareness that remained potentially accessible to consciousness or reflection, and, in any case, could not basically be foreign to the subject. In Husserl's words:

> What I do not 'know', what in my experience, my imagining, thinking, doing, is not present to me as perceived, remembered, thought, etc., will not 'influence' my mind. And what is not in my experience, be it ignored or implicitly-intentionally decided, does not motivate me even unconsciously
>
> (Husserl 1952: 231).

These two views seem hardly reconcilable. Antagonistic as they may seem, however, on closer analysis, psychoanalysis and phenomenology do in fact have a common starting point: it is in the Cartesian view of consciousness as 'clear and distinct perception', the assumption that consciousness is transparent to itself insofar as its own contents are concerned. For Husserl, the 'cogito' is the present evidence, the necessary 'appresentation' of all contents in the observing consciousness, without which they would melt or escape into the unreality of past or future. All memories, all ideas, all the possibilities of consciousness, must cling, as it were, to this evident present so as not to vanish. But Freud's view of consciousness is not much different: what is conscious is only '… the idea that is present in our consciousness and which we perceive' in each case (Freud 1943: 29). Consciousness, therefore, as in classical thought, is considered the space for current ideas or representations. The unconscious is then the space which is conceived as containing all the other ideas which are not present at a particular moment. Freud rejects an ambiguous knowing-unknowing consciousness for '… a consciousness of which one knows nothing seems to me many times more absurd than a psychic

unconscious' (Freud 1940b: 243). Consciousness must be transparent to itself or it is not consciousness at all.

Psychoanalysis thus rebelled against the classical philosophy of consciousness, and not only failed to overcome it but, without being aware of it, even adopted its premises. The situation is similar to that in today's conflict between neurobiology and classical philosophy: the sovereign, autonomous conscious subject that neurobiology believes it must dethrone is itself merely a dualistic construct. Separated from its body and its life, restricted to present 'mental states', the bodiless, and to this extent powerless, 'ego' becomes easy prey to neurobiological reductionism, and the role of the unconscious as the actually powerful substrate is now taken over by the material brain. With this, of course, there is the threat of naturalizing subjectivity in a way that could have a much more reifying effect than Freud's interpretation of man as *'homo natura'*, criticized by Binswanger (1957).

Now, the dimension of *embodiment of the subject* which was increasingly brought to the fore by phenomenology as time went on, could just as easily have become the core of psychoanalysis. Freud, as is well known, did not only see the origin of the ego in the body.[2] The body also played a decisive role in psychoanalytical drive theory, since this theory assumed a step-by-step development of partial drives, which are dominated by certain regions of the body, and whose 'destinies' permanently affect the development of the individual. Nevertheless, despite this concept, the dualism of body and mind made an impact on psychoanalytic theory. For Freud, in the final analysis, drives are not phenomena of the lived body, but objective-somatic quantities; and their representations do not belong to a libidinous body of the subject but are already part of the psyche as an inner, hidden apparatus where drive derivatives and drive energies are converted into one another and distributed to various levels of the psyche—an apparatus which can only be decoded on the basis of external signs such as body-language or by way of speech. In the end, the body thus remained interesting only as the seat of symbolic or imagined meanings, as a primary projection field for the psyche, so to speak, which always had to be scrutinized for its hidden meanings. That mental phenomena could at the same time be bodily as well was not imaginable in the dualistic paradigm.

With the idea of the 'psychic apparatus', which doubtlessly goes back to Freud's own early brain theory, an entity had also been created that served as a sort of inner container for images and memories of external reality. Introjected as 'object-representations', 'imagos' etc. they populated the various compartments of the psyche and there developed a life of their own with the help of the drive energies. In this way, the ego remained separated from important parts of these compartments through radical ignorance: the topologically structured, dynamic unconscious, according to Freud, is basically different from the pre-conscious as the latent and implicitly 'previously known' (Freud 1940c: 77f.). Between the preconscious and the unconscious stands the economical mechanism of repression, and both what is repressed and the repressing mechanism—i.e. the motivation for repression—elude consciousness. As evidence for this concept, Freud could point to physical symptoms or to *Freudian slips*, which appeared alien or meaningless to

the ego, furthermore to the difference between manifest and latent *dream content* which is attributable to an unconscious censor, and, last but not least, to the *resistance* shown by the patient during analysis to becoming aware of what has been repressed.

This radical separation of the unconscious, however, took place at the cost of its having to take a problematic ambiguous position between subjective experiencing and objective processes (Waldenfels 2002: 294). In fact, in the final analysis, it had to be assigned to the objectivity of the psychological apparatus. Freud solves the paradox he discovered, namely that one 'knows something that one simultaneously does not know' and that 'one is struck with blindness while the eyes see' (Freud 1957: 175n) by the splitting of the psyche into two parts. As a consequence, the unconscious turns into an 'internal foreign country' (Freud 1940c: 62), in other words to something external within oneself, whose meaning and effect are alien to the subject. At this point, however, one should not only bear in mind Husserl's objection to a motivation which is entirely alien to the subject. How, over and above this, should the subject be in the position to reappropriate such an alien meaning unless, both in origin and in its latency, it was always *his own* meaning? Psychoanalytical therapy could then do no more than convey rational insights into the mechanisms of one's own inner life, and could not contribute to a genuine integration of the personality. The aim of therapy: 'where id was, ego shall be', would then remain only a matter of explicit knowledge, not of actual self-appropriation.

The phenomenological critique of this concept now moved along various paths:

- Sartre saw the unconscious not as a circumstance imposing restrictions on the subject from outside, but as a basic modality of the subject's constitutive relationship to himself, namely, that of bad faith, '*mauvaise foi*' (Sartre 1958: 47ff). The subject assumes an ambivalent relationship to himself, he allows himself, so to speak, to slide into an 'intentional inattention': one doesn't know something *and* doesn't want to know it; one doesn't see something *and* doesn't want to see it, and in this way becomes the deceived and the deceiver in one.

- A comparable form of double consciousness may be found, as Bernet (1997) has undertaken to show, in Husserl's analyses of the perception of images, of the reproductive consciousness, of memory, and above all, of *imagination*: these forms of consciousness in each case entail a duplicity of presence and absence so that the ego lives in two worlds at the same time. In this way, they can also serve as paradigms for the relationship between conscious and unconscious.

- Another way of overcoming the dualism of conscious and unconscious consists of expanding the space of subjectivity *vertically* so to speak, so that it can include the phenomena of drive and urge as a basic stratum. This method of reinterpretation of Freud's metapsychological terms into an elementary activity of life which always precedes the conscious experience of oneself, was an approach partly adopted by Max Scheler (1928), and then, primarily, by Michel Henry (1992).

• Finally, there is the possibility of taking the ambiguity of the body, as understood by Merleau-Ponty, as the starting point, to extend subjectivity in the *horizontal* dimension and to encounter the unconscious in bodily behaviour, in day-to-day living, and in the structures of the person's lived space. *Body memory* plays a special part here, insofar as it changes a person's bodily and inter-bodily experiences into implicitly effective dispositions, which provide the mostly unconscious basis for day-to-day living (Fuchs 2000, 2012, 2016).

The latter is the course which I will take in what follows (without rejecting the other possibilities mentioned). So the question will be: can the unconscious be localized in the lived relationships and conduct of a person—in other words in the *horizontal* dimension of the lived body and intercorporeality? How far can such a concept reflect elements of Freud's unconscious? In what follows, I first want to develop the concept of body memory and the relational field that it constitutes, and then ask about the structures of this field where the unconscious can, as it were, take up its abode.

Body Memory

If, following Merleau-Ponty, we view the body not as the visible, touchable, and sentient physical body but first and foremost as our *capacity* to see, touch, and sense, then body memory designates the totality of these bodily predispositions as they have formed in the course of our development—in other words, in their historical dimension. In body memory, the situations and actions experienced in the past are, as it were, all fused together without any of them standing out individually. Through the repetition and superimposition of experiences, a habit structure has been formed: well-practiced motion sequences, repeatedly perceived *Gestalten*, forms of actions and interactions have become an implicit bodily knowledge and skill. Body memory does not take one back to the past, but conveys an implicit effectiveness of the past in the present. This approach converges with the results of recent memory research on the central significance of *implicit* memory which is just as much at the basis of our customary behaviour as of our unconscious *avoidance* of actions (Fuchs 2012; Schacter 1996).

The body is thus the ensemble of organically developed predispositions and capacities to perceive, to act, but also to desire and to communicate. Its experiences, anchored in body memory, blanket the environment like an invisible network which relates us to things and to people. It is, as Merleau-Ponty writes, 'our permanent means of "taking up attitudes" and thus constructing virtual presents', in other words to actualize our past and, with this, to make ourselves feel at home in situations (Merleau-Ponty 1962: 181). Even more: in the bodily experience structures, the other is always already included, he is understood in expression and intended in desire. Before I can reflect on what I am communicating through my gestures or speech, my body always already creates the feeling of being-with; it expresses itself through attitude and gestures, and at the same time reacts to the impressions of others. This 'intercorporeality' (Merleau-Ponty 1964: 168) forms an overriding, intersubjective system in which, from childhood on, forms of bodily interaction are established and constantly reactualized. It comprises the self and the others, the conscious and the unconscious: 'I do not have to search very far for others: I find them in my experience, lodged in the hollows that show what they see and what I fail to see (…) We are in no way locked inside ourselves' (Merleau-Ponty 1973: 138f).

Body Memory and Life Space

Body memory—like the body schema—thus forms not only an interior system restricted to the physical body. It also constitutes a sensorimotor, emotional, and interactive field in which we, as embodied beings, constantly move and conduct ourselves. The terminology of Kurt Lewin's field psychology (1936) offers itself here, particularly the concept of the *life space*. In order to link it with the structures of body memory, I want to give a brief outline of this in what follows.

The life space is centred around the person and the person's body. According to Lewin, it is characterized by experienced characteristics such as closeness or distance, narrowness or breadth, connectedness or separateness, attainability or elusiveness, and it is structured by physical or symbolic *boundaries* which offer resistance to movement. This produces more or less clearly bounded *sectors* such as the peripersonal space around one's own body, claimed territories (property, home), the sphere of influence which emanates from someone, but also prohibited or taboo zones. The lived space is further permeated by tangible '*field forces*' or '*vectors*', in the first place those which attract and repel. Competing attractive or repulsive forces in the life space lead to typical conflicts such as attraction *versus* aversion, attraction *versus* attraction etc. They can be considered as conflicting directions of movement or possibilities which are offered to a person in a given situation.

A good example of conflicting field forces is offered by the situation of a small child who is torn back and forth between his bond to his mother and curiosity (cf. Stern 1991: 101). The mother is first of all the 'safe haven', the centre of gravity, so to speak, which curves the child's experienced space in such a way that he remains in her vicinity. The space thus acquires a gradient: the further the child moves away from the mother, the more empty, more lonely the space becomes. While it condenses again around other, i.e. strange, people, the child rather makes a detour around them: the space curvature near them is 'negative'. Little by little, the child's exploratory drive and the attractive charms of the environment loosen his tie to his mother, so that it becomes possible to increase the distance against the gradient—only until the bond is stretched too much, and the child runs back to his mother in the end. This example is also a good illustration of the fact that the respective field structures are based on body memory, in this case, the history of the experiences the child has had in closeness and ties to his mother. Another proverbial example lies in the saying, 'Once bitten, twice shy', which illustrates the *aversive* effect of body memory. A third example, finally, is given by the zones of prohibition which restrict the directions in which a child can move so that his spontaneous impulses interfere with parental imperatives, namely, inasmuch these have left a negative mark on his very life space.

Consequently, the life space—depending on the respective experiences, capabilities, and motives of a person—can bear varying significances, relevances, or valences. In analogy to a physical field, 'gravitational effects', invisible 'curvatures' of space, or barriers can appear which restrict or prevent spontaneous movements. Particularly in psychopathology, we encounter various deformations of the lived space (Fuchs 2007), as, for instance, the taboo zones of obsessive patients and the avoidance zones of phobic patients, which are based on certain past experiences sedimented in body memory.

On the Phenomenology of the Unconscious

With this, I have made a brief sketch of an approach and a terminology which permit the question of the unconscious to be put and answered in a different way.

If we reject a topologically structured unconscious beyond consciousness—a separate intra-psychic process which impacts on the experiencing subject from outside, so to speak —then we may ask whether the unconscious might not be considered another mode of experiencing that manifests itself in the *horizontal* dimension of the lived body and the lived space. The paradigm for this would be the ambiguity of the body itself which, while seeing, always remains unseen, and of whose dispositions I often remain unaware, which in fact come to meet me from outside, namely in the form of the attractive or repelling objects, the affordances and field structures of my environment. Such an unconscious would then, as Merleau-Ponty writes, '… be sought not at the bottom of ourselves, behind the back of our "consciousness", but in front of us, as articulation of our field' (Merleau-Ponty 1968: 180). It would be the unrecognized reverse side of our experience and conduct, or its other, hidden meaning.

As our starting point, let us first consider the field structure of a repressed wish. In his short story '*Der Branntweinsäufer und die Berliner Glocken*' ('The brandy drinker and Berlin's bells'), Heinrich von Kleist recounts the story of an alcoholic soldier who, after insistent preaching and punishment, has resolved to become abstinent but was found drunk after only three days. Asked how this relapse could have happened after all his good resolutions, the soldier justified himself by saying that the devil must have had his hand in it because while walking through the town he suddenly heard the names of various brandies in the tolling of the bells—for example '*Kümmel! Kümmel!*' in the ringing of the town hall bell, '*Pomm*eranzen, *Pomm*eranzen' in the ringing cathedral bell, and so on. In the end, he could not help being defeated by these insidious sounds.[3] While this humorous example relates only to a wish that was not repressed but merely suppressed by an act of will, it gives a fine illustration of the indirect way in which contrary bodily impulses or drives can get their way, namely from outside. The experiential field is, so to speak, interspersed with suppressed desire which becomes crystallized finally around certain perceptions—namely those which are sufficiently vague while offering a certain similarity for the purpose: in this case the various chimes. The uncertain or *ambiguous* is the place where a latent or hidden significance can take shape. The drive or the wish that was not satisfied breaks through circuitously and from outside so that, in principle, we can already recognize the mechanism of *displacement*. What is actually desired is fulfilled through something similar.

A comparable interference of thematic and non-thematic (only bodily expressed) directions of meaning is also found in the various types of '*Freudian slips*'. Freud himself says that '… slips are the result of two different intentions which interfere with one another, of which one can be called the disturbed and the other the disturbing intention' (Freud 1940a: 56). Mishearing is most like the example of Kleist's soldier: a

latently desired meaning is 'interpreted' from a similar sequence of sounds. With mistakes in speaking, writing, and in (mis)placing things another intention interferes with the explicitly intended action, so that 'the right hand—literally—does not know what the left hand is doing'. Finally, with *forgetting*, an originally made but unpleasant intention is blanked out and replaced by others, for example, routine processes. Thus, in spontaneous bodily perceptions or actions which take place 'of their own accord', the relevant latent intention breaks through—in a reversal or interlacing which is linguistically expressed by the prefix 'mis-'.

The producer of the slip can now either immediately or after some brief thought recognize its significance and ascribe it to himself, or he finds it 'senseless', in other words, alien to himself. For example, Freud writes the following concerning 'misspeaking':

> If later we present it [the intention on which the misspeaking was based] to the speaker, he may either acknowledge it as something familiar, so that it was only temporarily unconscious, or he may deny it as alien to himself, which means that it was permanently unconscious
>
> (Freud 1940c: 77).

It is on this difference, amongst other things, that Freud bases his categorical distinction between the preconscious and the true dynamic unconscious which is excluded or repressed from consciousness 'by living forces' (Freud 1943: 436). The defence mechanism and the corresponding resistance to the latent meaning obviously have as their prerequisite that the inhibitive trends and their motives are themselves excluded from consciousness. However, the question is whether this justifies establishing a special intrapsychic space for the dynamic unconscious. Against this, there is the merely gradual difference between a temporary and a permanent unconscious in the Freud quotation just cited. In both cases, after all, we are dealing mainly with a duplicity of intentions, to which only an additional repressive tendency is added in the second case. But if we do not assign the 'living forces' of repression of which Freud speaks to an intrapsychic mechanism beyond consciousness, but see them rather as field forces, we will easily find models for them in the bodily or life space.

The first thing that comes to mind would be the *relieving posture* adopted after sustaining an injury: spontaneously one avoids putting the injured limb at risk from dangerous objects and holds it back without having still to think of the event. Avoidance behaviour is thus incorporated into the implicit body memory. Moreover, I have already mentioned the *zones of prohibition* which face the child and operate against its approach through negative field forces until the child respects them automatically. We come one step closer to the dynamic unconscious with zones or objects which are *taboo*. For, unlike prohibition, the taboo has a special structure and effect in that it is not expressly formulated but is generated by the avoidance behaviour of others, like a negative curvature of the shared

life space around what is prohibited. Taboos are most effective when the members of the community are not aware of them. The infringement of taboos is not necessarily punished with open penalties, but automatically generates feelings of shame, guilt, or abhorrence in the offender, reinforced by the contempt and the ostracizing silence of the others.

In all these cases, experience and conduct are determined by negative—i.e. 'repulsive'—field forces exercising their effect unconsciously since the subject, like the 'bitten' person, has gradually extricated herself from the possible conflict. Avoidance has become an implicit, bodily pattern of behaviour so that what is potentially threatening in the environment is no longer consciously experienced. Nevertheless, repelling forces do not appear to consciousness as coming from outside but, in Hegel's terms, as its own otherness. They remain coextensive with the experiential field but as its negative. The manifest feelings of fear, guilt, or shame which arise on stepping beyond the barriers in the life space were already latently present before, endowing these barriers with their affective loading.

In the same way, in the case of a 'slip', the dynamic unconscious puts up resistance to its becoming conscious. This resistance is itself not conscious, nor is it implicitly or preconscious, and yet on this account it is still not altogether outside consciousness. It is rather an *ambiguity or duplicity of consciousness itself*; in such a way that the subject, if she hits on the manifestation of the hidden meaning, at least has an inkling that it is asking her a *question*, namely about her own otherness. The unconscious, writes Merleau-Ponty:

> … cannot be a process 'in third person', since it is the unconscious which chooses what aspect of us will be admitted to official existence, which avoids the thoughts or situation we are resisting, and which is therefore not an *un-knowledge* but rather a non-recognized and unformulated knowledge that we do not want to take up. In an approximative language, Freud is on the point of discovering what other thinkers have more appropriately named *ambiguous perception*
>
> (Merleau-Ponty 2007: 194).

We can understand this ambiguity of consciousness with the example of another defence mechanism, namely *projection*. Here the beam in one's own eye becomes the splinter in another's eye; in other words, one perceives in others the impulses and motives against which one has built defences in oneself. Naturally, this perception is also ambiguous, since the excessive zeal with which the impulses in others are disapproved derives its energy precisely from the efforts one has to make to neutralize one's own impulses. The blind spot in self-awareness—and here Freud is doubtlessly right—does not result from a mere 'overlooking', but from active and emotionally charged repression. Nevertheless, this repression remains the work and the effort of the subject herself, not of a mechanism outside her.

Trauma and Reiteration

Let us now turn to another phenomenon, namely, the unconscious effect of an emotional trauma which Merleau-Ponty sets out to interpret in his 'Phenomenology of Perception'. What is repressed, he writes, is like a phantom limb for an amputee inasmuch as a bodily capacity continues in the latter which is no longer congruent with the present. Habitual and current body come into conflict with one another. Similarly, repression also creates an empty space in current subjectivity (Merleau-Ponty 1962: 87), as if the 'negative' left by an experience which has not been dealt with interposes itself unnoticed before every new situation and thus imprisons the traumatised person in a past which is still present.

> … (T)his fixation does not merge into memory; it even excludes memory in so far as the latter spreads out in front of us, like a picture, a former experience, whereas this past which remains our true present [the trauma, T.F.] does not leave us but remains constantly hidden behind our gaze instead being displayed before it. The traumatic experience does not survive as a representation in the mode of objective consciousness and as a 'dated' moment; it is of its essence to survive only as a manner of being and with a certain degree of generality
>
> (Merleau-Ponty 1962: 83).

This description assigns the repressed trauma not to explicit but to body memory, which holds what is hidden 'from sight' and goes on living in a general manner or 'style' of existence.[4] The injury has penetrated the body of the subject and has left behind a permanent responsiveness, a readiness to defend itself. The traumatized person becomes hypersensitive to threatening, shaming situations similar to the trauma in some manner, even if this similarity is not consciously known, and tries to circumvent them. 'The resistance is directed to a certain area of experience, a certain category, a certain type of memory' (Merleau-Ponty 1962: 194). All the same, at every step, the victim may encounter something that reawakens the trauma in her. Often it happens that a permanent predisposition develops to react with fear and nervousness, to become alarmed every time the doorbell rings, a feeling of being followed or observed by unknown people.

An impressive description is to be found in the memoirs of the Jewish writer Aharon Appelfeld, who from his seventh to his thirteenth year of age experienced the Second World War hiding in the woods of the Ukraine:

> More than fifty years have passed since the end of the war. I have forgotten much, even things that were very close to me—places in particular, dates, and the names of people—and yet I can still sense those days in every part of my body. Whenever it rains, it's cold, or a fierce wind is blowing, I am taken back to the ghetto, to the camp, or to the forests where I spent many days. Memory, it seems, has deep roots in the body. (…) The cells of my body apparently remember more than my mind

which is supposed to remember. For years after the war, I would walk neither in the middle of the sidewalk nor in the middle of the road. I always clung to the walls, always staying in the shade, and always walking rapidly, as if I were slipping away. (…). Sometimes, just the aroma of a certain dish, or the dampness of shoes or a sudden noise is enough to take me back in the middle of the war (…). The war has infiltrated my bones. (…) the palms of one's hands, the soles of one's feet, one's back, and one's knees remember more than memory. Had I drawn from them, I would have been overwhelmed with what I have seen

(Appelfeld 2009: 50, 90, vii).

Here it is not a particular episode, but an entire segment of his life that has left its mark on the body, more deeply and permanently, of course, than the autobiographic memory could have done: proprioception, touch, smell, hearing, even certain kinds of weather can suddenly allow the past to come to life again, and even bodily pattern of movement, such as the hunted walk close to the wall, still imitates the behaviour of the fugitive.

The effect of the trauma on the person can thus be viewed, first, as a specific deformation of her lived space corresponding to an unconscious avoidance behaviour which she adopts towards the anxiety-provoking or '*repelling zones*'. The lived space around these zones is to a certain extent negatively curved and prevents the free development of the life movement. Second, the life space is permeated with similarities in which the trauma approaches the traumatized person from outside, so that it is impossible to avoid it. For in one's attitude, one's stance, and in one's perceptive predispositions, one carries the trauma into one's world over and over again.

It is to this that the psychoanalytic concept of *repetition compulsion* relates. This is based on the clinical experience that patients continue to be drawn into the same, mostly damaging behaviour or relationship patterns even if they try to prevent this at the conscious level. Their lived space is so to speak 'positively curved' around these regions— in other words, these exercise an unnoticed *attraction*. If, for example, a person's early experiences were characterized by abusive and violent relationships, this issue will determine also that person's later relationship patterns. The types of abuse may vary, but the implicit behaviour patterns deposited in body memory will have the effect of fulfilling her expectations and bring about the familiar type of relationship. These unconscious enactments, as they are called today, were, of course, seen by Freud as a form of transference. As he writes, we must:

… say in analysis that the analysand *remembers* nothing at all of what has been forgotten and repressed, but he *acts it out*. He does not reproduce it as a memory but as action, he *repeats* it, naturally without realizing that he is repeating it. For example, the analysand does not say that he remembers being defiant and incredulous towards the authority of his parents, but he behaves in this manner towards the doctor

(Freud 1946: 129).

The unconscious prehistory of intersubjective relations is re-enacted through the intercorporeal memory. However, this means that the unconscious is not a hidden chamber of the psyche any more, but is interwoven in the lifestyle, in the bodily conduct of a person, as a substructure which remains hidden from her personally, but becomes visible to others because, in the final analysis, it is always implicitly directed to those others themselves. The 'blind spot' in the centre of consciousness can also be viewed as the other side of the intersubjective relationship, in which our own being-with-others must necessarily remain hidden from us, so that this dark side of ourselves can only be illuminated in our communication with others. For in my world they dwell '… in the hollows that show what they see and what I fail to see ' (Merleau-Ponty 1973: 138f).

Conclusion

From the point of view of a phenomenology of the lived body, the unconscious is not an intrapsychic reality residing in the depths 'below consciousness'. Rather, it surrounds and permeates conscious life, just as in picture puzzles the figure hidden in the background surrounds the foreground, and just as the lived body conceals itself while functioning. It is an unconscious which is not located in the *vertical* dimension of the psyche but rather in the *horizontal* dimension of lived space, most of all lodging in the intercorporeality of dealings with others, as the hidden reverse side of day-to-day living. It is an unconscious which is not to be found inside the individual but in his relationships to others.[5]

Unconscious fixations are like certain restrictions in a person's space of potentialities produced by an implicit but ever-present past which declines to take part in the continuing progress of life. Their traces, however, are not hidden in an inner psychic world but manifest themselves rather as 'blind spots', 'empty spaces', or curvatures in the lived space: in the 'slips' in speech and action; in the relationship patterns into which a person repeatedly blunders, in the actions which are avoided without being aware of it; in the spaces which are not entered, the opportunities offered by life which one does not take, and even does not dare to see. Such traces may be recognized as 'negatives' so to speak, in the form of inhibitions or omissions which are characteristic of a person. They can also become symbolically or physically present in neurotic or psychosomatic symptoms. The symptom is to this extent neither meaningless nor a defective habit—as learning theory assumed[6]—nor is its meaning to be found beyond itself, in the unconscious interior. Rather, it lies in the intercorporeal expression, in the interactive field, even if this meaning is not evident, but must be understood and interpreted.

The unconscious is thus absence in presence, the unperceived in the perceived (Merleau-Ponty 1968: 245f). Like a figure blanks out the ground from which it stands out, thus consciousness, perception, and language conceal their reverse side, namely the unconscious, the unperceived, and the silence, which are always bound up with them.

This reverse side, however, does not remain fully concealed but expresses itself in reversals, chiasmatic entanglements, in an ambiguity of consciousness: one does not know something *and* does not want to know it; one does not see something *and* does not want to see it—in other words, one looks past it intentionally–unintentionally. Consciousness is not fully transparent to itself because it hides itself from itself.

This duplicity of consciousness corresponds to the ambiguity of the body whose modes of appearing fluctuate between the thematic and the unthematic, between the physical (*Körper*) and the lived body (*Leib*). But it also corresponds to the ambivalent, conflict-prone nature of our existence itself where we, precisely as natural, embodied beings, can always confront our own instinctive and natural side as well. This is what constitutes the contradictoriness or, to speak with Plessner (1975), the 'eccentricity' of the way we relate to ourselves, the constant conflict between spontaneity and reflectivity, body and soul, nature and nurture, conscious and unconscious. One could then accuse Freud that even he, for all his scepticism, was too generous to humankind in that he tried to relieve our consciousness of this inherent conflict by placing our opposing will in a separate space belonging to the unconscious—thus withdrawing this will from our responsibility.[7]

References

Appelfeld, A. (2009). *The Story of a Life: A Memoir*. New York: Random House.

Bernet, R. (1997). 'Husserls Begriff des Phantasiebewusstseins als Fundierung von Freuds Begriff des Unbewussten'. In C. Jamme, *Grundlinien der Vernunftkritik* (S. 277–306). Frankfurt/M: Suhrkamp.

Binswanger, L. (1957). 'Freuds Auffassung des Menschen im Lichte der Anthropologie'. In Ders., *Der Mensch in der Psychiatrie*. Pfullingen: Neske.

Eysenck H. J. and Rachmann, S. (1972). *Neurosen—Ursachen and Heilmethoden*. Berlin: Deutscher Verlage der Wissenschaften.

Freud, S. (1940a). *Vorlesungen zur Einführung in die Psychoanalyse*. GW Bd. 11. Frankfurt: Fischer.

Freud, S. (1940b). *Das Ich and das Es*. GW Bd. 13, S. 235–289. Frankfurt: Fischer.

Freud, S. (1940c). *Neue Folge der Vorlesungen zur Einführung in die psychoanalysis*. GW Bd. 15. Frankfurt: Fischer.

Freud, S. (1943). *Einige Bemerkungen über den Begriff des Unbewussten in der psychoanalysis*. GW Bd. 8, S. 429–439. Frankfurt: Fischer.

Freud, S. (1946). *Erinnern, Wiederholen, Durcharbeiten*. GW Bd. 10, S. 126–136. Frankfurt: Fischer.

Freud, S. (1957). *Studien über Hysterie* (mit J. Breuer). GW Bd. 1, S. 77–312. Frankfurt: Fischer 1952.

Fuchs, T. (2000). 'Das Gedächtnis des Leibes'. *Phänomenologische Forschungen 5*: 71–89.

Fuchs, T. (2007). 'Psychotherapy of the lived space. A phenomenological and ecological concept'. *American Journal of Psychotherapy 61*: 432–439.

Fuchs, T. (2012). 'The phenomenology of body memory'. In S. Koch, T. Fuchs, M. Summa, and C. Müller (eds), *Body Memory, Metaphor and Movement* (pp. 9–22). Amsterdam: John Benjamins.

Fuchs, T. (2016). 'Embodied knowledge—embodied memory'. In S. Rinofner-Kreidl and H. Wiltsche (eds). *Analytic and Continental Philosophy. Methods and Perspectives* (pp. 215–229). Berlin: De Gruyter.

Graumann, C. F. (1960). *Grundlagen einer Phänomenologie und Psychologie der Perspektivität*. Berlin: De Gruyter.

Henry, M. (1992). *Radikale Lebensphänomenologie. Ausgewählte Studien zur Phänomenologie*, hrsg. R. Kühn. Freiburg München: Alber.

Husserl, E. (1952). *Ideen zu einer reinen Phänomenologie and phänomenologischen Philosophie. II. Phänomenologische Untersuchungen zur Konstitution*. Den Haag: Nijhoff.

Kleist, H. v. (1984). *Sämtliche Erzählungen and andere Prosa*. Stuttgart: Reclam.

Lewin, K. (1936). *Principles of topological psychology*. New York: McGraw-Hill.

Merleau-Ponty, M. (1962). *Phenomenology of perception*, trans. Colin Smith. London: Routledge & Kegan Paul (French original: *Phénomenologie de la perception*. Paris: Gallimard, 1945).

Merleau-Ponty, M. (1964). 'The philosopher and his shadow'. In *Signs*, trans. A. McClaery (pp. 159–181). Evanston, IL: Northwestern University Press.

Merleau-Ponty, M. (1968). *The visible and the invisible*, trans. A. Lingis. Evanston, IL: Northwestern University Press.

Merleau-Ponty, M. (1973). *Adventures of the dialectic*, trans. J. Bien. Evanston, IL: Northwestern University Press.

Merleau-Ponty, M. (2007). 'Man and adversity'. In T. Toadvine and L. Lawlor (eds), *The Merleau-Ponty Reader* (pp. 189–240). Evanston, IL: Northwestern University Press.

Plessner, H. (1975). *Die Stufen des Organischen and der Mensch*. Berlin: De Gruyter.

Romanyshyn, R. D. (1977). 'Phenomenology and Psychoanalysis'. *Psychoanalytic Review 64*: 211–223.

Sartre, J.-P. (1958/1969). *Being and Nothingness: An essay on phenomenological ontology*, trans. H. E. Barnes. London: Methuen and Co.

Schacter, D. L. (1996). *Searching for memory: The brain, the mind, and the past*. New York: Basic Books.

Scheler, M. (1928). *Die Stellung des Menschen im Kosmos*. Darmstadt: Reichl.

Stern, D. N. (1991). *Tagebuch eines Babys. Was ein Kind sieht, spürt, fühlt and denkt*. München: Piper.

Waldenfels, B. (2002). *Bruchlinien der Erfahrung; Phänomenologie—Psychoanalyse—Phänomenotechnik*. Frankfurt/M.: Suhrkamp.

Notes:

(1.) As is well known, both Husserl's '*Logische Untersuchungen*' and Freud's '*Traumdeutung*' appeared in 1900.

(2.) Cf. Freud 1940b: 253.

(3.) Kleist 1984. The story is also cited by Graumann (1960: 151) as an illustration of the motivational basis of perspectivity.

(4.) 'comme un style d'être', in the French original (Merleau-Ponty 1945: 98).

(5.) '(…) the latency of psychoanalysis is an unconscious that is *beneath* conscious life and *within* the individual, an *intrapsychic* reality that leads to a psychology of depth in the *vertical* dimension. (…) the latency of phenomenology is an unconscious which *surrounds* conscious life, an unconsciousness in the world, *between us*, an *ontological* theme that leads to a psychology of depth in the *lateral* dimension' (Romanyshyn 1977).

(6.) 'Learning theory assumes no "unconscious" causes whatsoever but views neurotic symptoms simply as learned habits. There is no neurosis at the bottom of the symptom, only the symptom itself ' (Eysenck and Rachmann 1972: 20).

(7.) Acknowledgement: Originally published as Fuchs, Thomas (2012), 'Body Memory and the Unconscious', in Dieter Lohmar and Jagna Brudzinska (eds), *Founding Psychoanalysis Phenomenologically: Phenomenological Theory of Subjectivity and the Psychoanalytical Experience* (pp. 69–82). © Springer Science+Business Media B.V.

Thomas Fuchs

Karl Jaspers Professor of Philosophy and Psychiatry Psychiatric Department, University of Heidelberg

Oxford Handbooks Online

Abstract and Keywords

This introduction provides an overview of the chapters in this section, which explores the role of psychoanalysis in aesthetics. More specifically, the chapters examine some psychoanalytic concepts with which to think more deeply about human creativity and aesthetic sensibility, such as wish and wish fulfilment, the depressive position, projection, containment, and mentalization. The focus is on what Sigmund Freud thinks about art, how we should understand it (the question of criticism), what makes an experience distinctively aesthetic, and how we should understand artistic creativity. One of the chapters deals with film theory, arguing against the cognitive turn in favour of the view that 'the creation and experience of film is driven by desire and wish fulfilment and functions so as to satisfy certain psychological, protective, expressive needs of both artists and audiences'. Another chapter considers the developmental, transformative nature of art, and the particular importance of its form in this respect.

Keywords: psychoanalysis, aesthetics, wish-fulfilment, projection, containment, mentalization, Sigmund Freud, art, artistic creativity, film theory

In our 'Introduction: Know Thyself', we commented that one important use that philosophers have made of psychoanalysis is to deepen our understanding of the meaning and meaningfulness of human activity and experience. Among the defining characteristics of human beings are our creativity and aesthetic sensibility. Experiences of creativity and aesthetic appreciation, which can be unusually intense, are essential to human life, and yet very difficult to articulate and understand reflectively. As Freeman notes in his chapter on the work of Richard Wollheim, arguably the most significant philosopher of psychoanalysis of the twentieth century, philosophical aesthetics seeks to think and write about these experiences in a rigorous and analytic way, and psychoanalysis can assist in that.

Aesthetics: Introduction

We may take the following as core questions that the philosophy of art seeks to answer: What is art? How should we understand it (the question of criticism)? What makes an experience distinctively aesthetic? And how should we understand artistic creativity? On each of these, Freud has little to say directly, and although he attempts to address the second and fourth (see Freud 1924 and 1908 respectively), Wollheim (1970) argues that he does not display a deep understanding of art. But his initial thoughts point us in particular directions that may be fruitfully developed.

Freud (1908) famously claims that art, like dreaming, involves the creation of a fantasy in which unconscious wishes are fulfilled. As in dreaming, we enjoy these wish-fulfilling fantasies in many cases because the wish is at least partly disguised. Unlike dreaming, however, to engage the spectator, not merely the artist, the fantasy must touch on something universal in the human condition. The formalist Roger Fry (1924) rightly objected that art cannot be understood simply as wish-fulfilment, but many of his objections rest on misunderstanding Freud's claims or can be met by a more sophisticated account of the place of wish-fulfilment in artistic creativity. For example, Fry overlooks Freud's equal emphasis of the importance in art of both taking account of reality (e.g. the manipulation of real materials) and connecting the wish to something universally true in human experience. Or again, many emotions and wishes may be at play, and the wish that is fulfilled is to find an expression that represents a resolution of the conflict between them.

Building on this idea of conflict resolution, Hanna Segal (1985, 1991) famously develops a detailed account of artistic creativity through the lens of Kleinian theory. Leaving aside for now Klein's backstory of distinct developmental *stages*, the centrepiece of the theory as it has been taken up in post-Kleinian work is that of discrete psychological *positions*. The depressive position is that mode in which we are able to experience both love and hate towards the same object, giving rise to anxiety, feelings of loss and guilt, and the desire to 'repair' symbolically what we destroy in feeling. The impulse to create arises from such feelings, to resolve the (usually unconscious) conflict of the depressive position. Segal famously provides an analysis of classical tragedy in these terms, in which she drew on Bion's notion of the container–contained to articulate the disturbing content of the play as contained—made bearable and thus able to be thought about—by the beautiful, unifying form it takes. A related Kleinian account is provided by the art critic Adrian Stokes (2014) who understood the total form of a painting or sculpture as the container for that turmoil projected into it by the artist who can there work it through into more manageable and harmonious forms.

Thus, psychoanalytic theory, as Freeman notes, provides the grounds for defending the 'psychologization'—or given the anti-psychologization of the last seventy years following Wimsatt's and Beardsley's (1946) famous attack on 'the intentional fallacy', a 'repsychologization'—of the meaning of an artwork, one which connects it closely to why we produce art at all and why we seek it out. Wollheim's response to Wimsatt's and Beardsley's charge is carefully laid out in 'Criticism as Retrieval' (1980), but each chapter

included here provides additional reasons to revisit psychoanalytic approaches in aesthetics.

It is worth noting that the rejection of psychoanalytic contributions to understanding art has been fuelled by two misunderstandings that mark discussions from Fry onwards.[1] The first, already noted, is a misunderstanding of psychoanalytic theory itself, typically substituting its rich resources for an implausible oversimplification. The second is the thought that psychoanalytic accounts of criticism or creativity aim to be complete in themselves; because they are not, psychoanalytic approaches are to be rejected. But while it is true that psychoanalytic resources must be supplemented by others, this does not make them irrelevant. As Cox and Levine note in their chapter, the light that psychoanalysis casts on our experience of art can be necessary without being sufficient. And Wollheim (1980) would agree, arguing that there is much else that we need to know as well, from artistic conventions to available materials to the social purposes of artworks.

The importance of desire in aesthetics forms the main theme of Cox's and Levine's chapter on film theory. Arguing against the cognitive turn, perhaps most forcefully championed by Noel Carroll, they defend the view that 'the creation and experience of film is driven by desire and wish-fulfilment and functions so as to satisfy certain psychological, protective, expressive needs of both artists and audiences'. Without this perspective, we cannot adequately explain the power of film and its effects on viewers, nor some characteristics of specific genres, such as horror.

It is notable that Wollheim himself goes well beyond wish-fulfilment in tapping the resources psychoanalysis provides. In criticism, says Wollheim, we should seek to reconstruct the creative process, but this includes not just conscious and unconscious intentions, but many things—psychological, historical, social, cultural, and chance—that can impact on these. While Freud himself does not make the connection, Wollheim takes projection to be central to the processes of creation and appreciation. Our imagination redeploys what has its psychological origins as a defence. Freeman provides an admirably clear guide to Wollheim's complex and challenging theory of how projection underpins our ability to perceive artworks and nature as expressive of psychological—especially emotional—properties. The role and nature of expressive perception forms the basis of Wollheim's account of what is both distinctive and valuable about art.

Galgut's discussion of literary form returns us to Segal's ideas of another value in art, namely its developmental, transformative nature, and the particular importance of its form in this respect. Many have noted that art resembles psychoanalysis in enabling emotions to be worked through, mostly notably in the case of literary art which uses language to give form to what we feel unconsciously. As we engage with reading, our minds relate to other minds, and this provides an opportunity for development, especially in emotional insight. The development of such insight—the ability to understand the emotional experience of others and oneself—is of course central to psychoanalysis but, Galgut argues, it can also be facilitated by literature. In philosophical and psychological

discussions of understanding other people, subsumed in debates over 'theory of mind', this attention to emotion—including the impact of emotional self-regulation on interpretation—is typically missing. By contrast, it plays a central role in the theory of 'mentalization', a modern development of a psychoanalytic account of how we understand ourselves and others (Fonagy et al. 2002). Galgut argues that certain kinds of literature, especially in respect of their form, can develop mentalization, not least through the complex provision of a multiplicity of perspectives.

As noted at the outset, then, psychoanalysis offers us a number of concepts—we have considered here wish and wish-fulfilment, the depressive position, projection, containment, and mentalization—with which to think more deeply about an aspect of human experience that is important, intensely felt, but otherwise enormously difficult to articulate and understand.

References

Fonagy, P., Gergely, G., Jurist, E., and Target, M. (2002). *Affect Regulation, Mentalization, and the Development of the Self*. New York: Other Press.

Freud, S. (1908). 'Creative writers and day-dreaming'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 9 (pp. 143–153). London: Hogarth Press/Institute of Psychoanalysis.

Freud, S. (1924). 'The Moses of Michelangelo'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 13 (pp. 211–238). London: Hogarth Press/Institute of Psychoanalysis.

Fry, R. (1924). *The artist and psychoanalysis*. London: Hogarth Press.

Segal, H. (1985). 'A Psycho-Analytical Approach to Aesthetics'. In P. Heimann and R. Money-Kyrle (eds), *New Directions in Psycho-Analysis* (pp. 384–406). London: Karnac.

Segal, H. (1991). *Dream, Phantasy and Art*. Hove/New York: Brunner-Routledge.

Stokes, A. (2014). *Art and Analysis*: *An Adrian Stokes Reader*, ed. M. Harris Williams. London: Karnac Books.

Wimsatt, W. and Beardsley, M (1946). 'The intentional fallacy'. *The Sewanee Review 54*: 468–488.

Wollheim, R. (1970). 'Freud and the Understanding of Art'. *British Journal of Aesthetics 10*: 211–224.

Wollheim, R. (1980). 'Criticism as Retrieval'. In *Art and its Objects: With Six Supplementary Essays* (pp. 185–204). Cambridge: Cambridge University Press.

Notes:

(1) Of course, in addition, if psychoanalysis provides a false theory of the human mind, it cannot correctly illuminate human aesthetic experience. We cannot here address the question of the truth of psychoanalysis, though we note in passing that the question is not simple—many psychoanalytic claims about mental functioning are most likely false while others are very well supported and most likely true—and it is unfortunate when psychoanalysis as a whole is rejected as a result of its misrepresentation.

Michael Lacewing

Michael Lacewing, Department of Philosophy, Heythrop College, University of London, London, UK

Richard G.T. Gipps

Richard G.T. Gipps, Oxford University, Disability Advisory Service, Oxford, UK

Oxford Handbooks Online

Abstract and Keywords

This chapter discusses the way in which Wollheim recruits the insights of psychoanalysis to develop a distinctive approach to philosophy of art in particular, and other branches of philosophy more generally, such as philosophy of mind and ethics. Wollheim develops an original approach to artistic expression and expressive perception of emotion by drawing on psychoanalysts' treatment of projection. He also develops an original approach to art criticism which he dubs criticism as retrieval, and which involves what he dubs the 'repsychologization' of pictorial meaning.

Keywords: art and emotion, expression, art criticism, Richard Wollheim, philosophy of art, repsychologization

Damien Freeman

(p. 477) On an Analytical Philosopher of a Psychoanalytical Bent

WRITING in the *Guardian* on 5 November 2003, Arthur Danto began his obituary for Richard Wollheim thus:

> The philosopher Richard Wollheim, who has died aged 80, belonged in the top echelon of thinkers who redefined the practice of his subject in Britain and the United States after the second world war. In terms both of the clarity of his writing and the acuity and ingenuity of his arguments, he embodied the intellectual virtues of analytical philosophy. But in terms of what engaged him as a philosopher, he stood far closer than any of his peers to continental thought.

On Richard Wollheim's Psychoanalytically Informed Philosophy of Art

This is the riddle of Wollheim's philosophy: he exemplified the virtues of analytical philosophy, but he was drawn to the preoccupations of the continental philosophers. Danto continues his obituary by explaining:

> Wollheim had little interest in donnish preoccupations with linguistic usage, or with the endlessly agonising issues of how language relates to reality. But he freely adapted some of the strategies worked out in addressing these issues to the problems that did engross him, which typically derived less from what other philosophers said than from what was central in his life.

(p. 478) What was central to his life might broadly be described as the 'irrational' aspects of human experience. He was captivated by the intensity of certain experiences and the grip that they have on the way we live our lives in virtue of their intensity. Chief among these were the experiences offered by art, in particular, painting as an art. Where modern philosophy in the Western tradition has been preoccupied with rationality, Wollheim applied the techniques of philosophy to speculate about the heart and the particular thought processes that govern it. Indeed, at the end of his obituary, Danto remarks:

> The heart was really the focus of his thought, in life and in philosophy, and it was the heart, above all, that he sought in the painting about which he was so passionate. The heart has not been the favoured organ of philosophical interest since perhaps Pascal, and it is this that set Wollheim apart from his peers in a discipline to which he brought originality and distinction.

Central to Wollheim's philosophical achievement was the idiosyncratic way in which he recruited the hypotheses of psychoanalysis as an aid to philosophical speculation. Wollheim was an expert on psychoanalytical theory, and he published numerous works discussing the philosophical issues surrounding psychoanalytical theory. However, he also did something more than that. His command of psychoanalytical theory and philosophy was such that he was able to use psychoanalysis to extend the way that philosophy speculates about human experience. It was this unique application of psychoanalysis that enabled Wollheim to apply the approach of analytical philosophers so effectively to the subject matter of continental philosophy. In this way, he might be said to develop a *psychoanalytically informed philosophy* rather than a philosophy of psychoanalysis.

This chapter seeks to introduce Wollheim's use of psychoanalysis in addressing the problems of philosophical aesthetics, and the rewards of his approach, by considering his relationship with art as practitioner, critic, and theorist, before saying something about his interest in psychoanalysis, and then investigating how art and psychoanalysis come together in his philosophy.

As discussed later, one feature that stands out here is Wollheim's treatment of projection. Projection has a venerable history both in philosophy and psychoanalysis. But Wollheim offers a unique—if, at times, controversial—approach to projection. In his philosophy of art, projection is central to his account of expressive properties that works of art possess, or, more specifically, to his account of our capacity for expressive perception of art and

nature. He offers a psychoanalytically informed account of projection, and this enables him to provide a new solution to an old problem of aesthetics. His account is not without its problems. What is fascinating, however, is the way it resonates with the intensity of his experience of works of art, and the astonishing readings of masterpieces that this permits.

His psychoanalytically informed approach to art also allows him to revisit the classic problem of the nature of art criticism. Where it had once been conventional to think about the critic's role as identifying the artist's intention in creating the work, (p. 479) Wollheim argues that we do better to think of it as retrieving the artistic process, and this project of criticism pursued as retrieval turns out to bear a strong resemblance to psychoanalysis, although he distances himself from claiming that it amounts to psychoanalysis.

In assessing Wollheim's legacy, we do well to consider what he exemplifies for the budding philosopher, as much as his place in the canon. For, while it is true that he has largely been marginalized, if not neglected, by mainstream analytical philosophy, he remains a truly inspiring figure for those who, like him, are consumed by the depths of human experience, and the need to provide some form of speculative account of these; an account worthy of their place in our lives.[1]

On Art and the Life of the Man

Art was central to Wollheim's life. In order to understand his philosophy of art, one has to see it in the context of his life.

It would not be in order to say that Wollheim was first and foremost an artist, but he did try his hand at it. In 1969, he published a novel, *A Family Romance* (Wollheim 1969). The semi-autobigraphical novel takes Freud's 'family romance'—a phantasy in which a child imagines that he is able to overcome parental authority—as the basis for a story about a man who executes a plan to poison his wife. Wollheim's novel is significant for the way it brings together his personal life, his interest in art, and his knowledge of psychoanalysis. Towards the end of his life, he completed *Germs* (Wollheim 2004), a memoir of childhood, which was published posthumously. Again, the memoir is significant for the way in which it draws together Wollheim's Proustian literary style, his reflections on his early life, and the self-analysis that his knowledge of psychoanalysis brings to these.[2]

Wollheim attracted considerably more acclaim for his writing as an art critic. He was a regular contributor to the contemporary art magazine, *Modern Painters*, and an acknowledged expert on the Western tradition of painting. That he was a connoisseur in the traditional sense is attested by invitations to address institutions devoted to such an approach, including the Department of History of Art at the University of Cambridge, where he presented 'Giovanni Morelli and the Origins of Scientific Connoisseurship', reproduced in *On Art and the Mind* (Wollheim 1974). Wollheim had a singular capacity for giving sustained attention to a painting, and apparently could sit in front of a single

painting for hours. Such sustained attention enabled him to attend to the details of the work, and, through these details, to gain a deep understanding of what they reveal about the meaning of the work of art. In this sense, he was a true connoisseur. But it is also (p. 480) interesting to consider the similarity between the connoisseur's attention to details and what they reveal on the one hand, and the psychoanalyst's attention to what seemingly trivial details reveal on the other. His art criticism blends together with his philosophy of art (and his application of the hypotheses of psychoanalysis) in the most celebrated achievement of his mature philosophy: the Mellon Lectures that he gave at the National Gallery of Art in Washington in 1984, subsequently published as *Painting as an Art* (Wollheim 1987).

Adrian Stokes (1902–72) was an artist and writer who exerted a profound personal and intellectual influence on Wollheim, a debt repaid in the form of a collection of Stokes's essays that Wollheim edited (Stokes 1972). Stokes was psychoanalysed by Melanie Klein in the 1930s, and this experience deeply informed his art criticism. In his introduction (Wollheim 1972) to *The Image in Form: Selected Writings of Adrian Stokes*, Wollheim discusses the influence of Kleinian psychoanalytic theory on Stokes's criticism, notably Stokes's treatment of the distinction between carving and modelling in sculpture in terms of Klein's paranoid-schizoid and depressive positions. Wollheim's introduction and Stokes's essays are of interest both for the intrinsic value of Stokes's psychoanalytically informed criticism and for the influence that this has on Wollheim's psychoanalytically informed philosophy. Wollheim explains of Stokes (1972: 29):

> by the time we reach the later writings, it becomes for the reader the most natural thing in the world—in the world of art, at any rate—to assimilate a shape to a feeling, or to equate the use of a specific material with a fantasy. In these works we catch the unmistakable voice of psychoanalytic culture.

This voice is likewise a feature of Wollheim's late writings, where he develops similar assimilations.

However, Wollheim was first and foremost a philosopher of art. Whereas 'epistemology' and 'theory of knowledge' are interchangeable, 'aesthetics' and 'philosophy of art' are not necessarily so. Aesthetics in modern philosophy is traditionally dated to Baumgarten's (1750) *Aesthetica*, where it means knowledge obtained through the senses, and then to Kant's (1790/1987) treatment of aesthetics as the study of the pleasure we derive from sensory experience of natural objects and artefacts. Under the influence of Hegel (1835/1975), however, there was a shift away from thinking about the faculty of taste and the ability to appreciate natural beauty, towards thinking about works of art as manifestations of the spirit of the age. This resulted in a preoccupation with art as a cultural practice that leads Hepburn (1966) to lament philosophy's neglect of natural beauty. This fixation with art as a distinctive cultural tradition gave rise to Duchamp's *Fountain* and Warhol's *Brillo Boxes*, and to the institutional theories of art, formulated and popularized by Danto (1964), and embellished with greater detail by Dickie (1974), according to which any object could count as a work of art, irrespective of whether it had

any inherently valuable properties, just so long as it was accorded the status of a work of art by a representative of the art world.

(p. 481) Wollheim was a staunch critic of the institutional theory of art, and published an influential essay refuting it, 'The institutional theory of art' (Wollheim 1980). He was, however, much more attuned to the philosophy of art that gave rise to this theory than he was to the older idea of aesthetics. That he privileges the artistic over the aesthetic is seen most strikingly in his account of the aesthetic attitude. Despite a long tradition identifying the aesthetic attitude with disinterested pleasure (Shaftsbury 1900) or an experience involving 'psychical distance' (Bullough 1912), Wollheim's solution is to assert that the aesthetic attitude is nothing more than the mode in which one attends to an object as a work of art (whether or not the object is, in fact, a work of art).

Art and Its Objects has received much critical acclaim as a systematic statement of the problems of the philosophy of art. It poses the question 'What is art?' at the outset, and introduces a working hypothesis that, in all art forms, a work of art is a physical object. The investigation of this problem forms the basis for a wide-ranging essay. While there are some indications of the influence of Freud, these tend to be passing references scattered throughout.

It is in *Painting as an Art* that Freud's influence is most keenly seen. The book is a study of the philosophy of painting, discussing what the artist does and what the spectator sees when looking at a painting executed by a genuine artist. In developing his arguments, however, Wollheim embarks on indepth discussions of specific paintings. These are much more than mere examples to illustrate philosophical points. They are highly sophisticated readings of paintings, in which all Wollheim's prowess as art historian and critic is on display. Psychoanalysis informs his theoretical account of the capacities that spectators draw on when looking at paintings, and the use artists make of these, as well as his reading of specific paintings, to extraordinary effect.[3]

On Wollheim's Philosophy and his Psychoanalysis

In an endorsement printed on the back cover of *The Thread of Life* (Wollheim 1984), Bernard Williams is quoted as summing up Wollheim's philosophy in the following terms:

> Wollheim's is a very original and extremely interesting approach to the philosophy of the mind, bringing together three areas of reflection not often combined: the philosophy of mind, psychoanalysis, and ethics.

That is a good statement both of what is original about his approach, and what is valuable about it. Wollheim's subject matter is difficult to pin down in the conventional (p. 482) terms of analytical philosophy. He was interested in practical life and the psychology that informs how we live our lives. More particularly, he was interested in the phenomenology of mental states—a peculiar combination of the intentionality and the feeling of the

various mental states—and the way in which the intensity of the feeling of certain mental states affects the way we live our lives. This may be found in two volumes of collected essays—*On Art and the Mind* and *The Mind and its Depths* (Wollheim 1993)—as well as two monographs—*The Thread of Life* and *On the Emotions* (Wollheim 1999).

The Thread of Life is a revised version of the William James Lectures that Wollheim gave at Harvard University in 1982. The purported subject of the lectures is the question: 'What is it to live the life of a person?' In providing an answer to this question, Wollheim touches on a range of issues including personal identity, the nature of the mind, and meta-ethics. To each of these, he brings a psychoanalytical perspective, and, although he demonstrates an understanding of the contemporary literature on these topics, his aim is not to address them head-on, but to develop his own account of what it is to live the life of the person, while showing, along the way, how his theory addresses contemporary issues. Drawing on psychoanalysis, Wollheim sketches out an account in which living the life of a person is a process in which a person's past manifests itself (in a disguised way) in the person's psychology such that it affects the person's mental states, dispositions, and behaviour by colouring the person's present orientation to the future through a distorting lens of the past.[4]

On the Emotions is a revised version of the Ernst Cassirer Lectures that Wollheim gave at Yale University in 1991. It provides a philosophy of emotion that is addressed only in passing in *The Thread of Life*. Wollheim does not give a conceptual analysis of emotion, but instead argues that the emotions must be understood in terms of their characteristic history, through which a past experience of satisfaction or frustration of desire comes to colour the way we perceive the object that we believe precipitates this satisfaction or frustration. The emotion is the attitude that we develop towards the precipitating factor, when our attention shifts from the earlier experience of satisfaction or frustration and settles on the object that precipitated it. We then experience this as the object of an attitude (or emotion) that colours our perception of the object in light of the past experience.

With this spotty précis of Wollheim's philosophy, our attention turns to his approach to psychoanalysis. It is necessary to say something about his experience and understanding of psychoanalysis, before considering how this fits into the general approach of his mature philosophy. Wollheim was not a psychoanalyst. He did undergo a lengthy analysis with the Kleinian psychoanalyst Dr Leslie Sohn. His analysis very clearly is an implicit influence on much of his writing, and is sometimes discussed explicitly.[5]

(p. 483) Wollheim's most sustained treatment of psychoanalysis is to be found in his *Freud* (1991). As an intellectual biography of Sigmund Freud it strives to afford an exposition of Freud's thought and its development. It stands in contrast to Wollheim's later (and more original) work, in which he develops ideas that he finds in Freud (and Melanie Klein), rather than providing a faithful exposition of them.[6] Given Wollheim's interests, two observations about *Freud* are particularly pertinent: the first concerns Wollheim's

conviction that Freud should be understood as a scientist and psychoanalysis as a science; the second, that Freud does not offer much insight into art.

In *Freud*, Wollheim is adamant that Freud's theory of mind and psychology needs to be seen as a science that is continuous with neurology, and, to this end, devotes the second chapter to the aborted *Project for a Scientific Psychology* (Freud 1895). Wollheim subsequently defends Freud's empirical method in 'Desire, Belief, and Professor Grünbaum's Freud' (in Wollheim 1993). In the last chapter of *Freud*, he deals with fragments of Freud's social thought, but acknowledges that there is nothing approaching a systematic social theory or social ethic. He says nothing about Freud's comments on art, although he does deal with Freud's remarks about Leonardo, Michaelangelo's *Moses*, etc. in 'Freud and the Understanding of Art' (in Wollheim 1974).

These points are worth emphasizing because they might seem to be at odds with Wollheim's project. He is interested in art, culture, psychology, and reflecting on the human condition. He believes that psychoanalysis has much to offer our understanding of these things. Yet he is also convinced that psychoanalysis is properly to be understood as a science, and that Freud, despite his personal interest in art, does not demonstrate a profound understanding of it in his writing. Of course one might take exception to these claims. What is interesting, however, is the way he is able to recruit the insights of Freud and psychoanalysis notwithstanding his convictions about them.[7]

Wollheim believes that psychoanalytic theory provides an accurate account of human psychology. And he is quite clear that the most accurate psychoanalytical theory is the one that descends from Freud to Melanie Klein and thence the Kleinian school (which included his own analyst, Leslie Sohn), following the British Psychoanalytical Society's 'controversial discussions' of 1942–4. It is not possible to investigate comprehensively (p. 484) his treatment of psychoanalysis in this chapter. Instead, attention shall focus on a single idea that is central to classical psychoanalytical theory and the psychoanalytically informed philosophy that Wollheim develops by drawing on it: projection.[8]

On Projection and the Human Condition

Although projection is one of the most celebrated concepts of the psychoanalytic corpus, it also has a history in philosophy that stretches back well before Freud. Freud introduced projection into psychoanalysis as a defence mechanism (e.g. Freud 1911). It is a means of coping with anxiety or a painful emotion: the mind projects its own mental state onto another mind, and, in doing so, gains a measure of relief. In philosophy, it can be traced back at least as far as Hume (1739–40/2000). For Hume, projection is important for understanding the metaphysics of fact and value. Simon Blackburn (1985) argues that projection is important for understanding the way we use language when we talk about value.[9] Both are doing something quite different from Freud, for whom projection is important for understanding the psychology of coping with anxiety and painful emotions.

Wollheim is interested in Freud's sense of projection as a defence mechanism, but we can also see in his work a concern with the philosophical categories of belief, evaluation, and perception. So he connects up the psychological and philosophical investigations. When Wollheim discusses 'projection', he is talking about a capacity of the mind to extend something mental into the world outside the mind. The mind, he suggests, owes this capacity to the Freudian defence mechanism. Wollheim's concept of projection is best understood as a redeployment of a capacity that belongs to a Freudian defence mechanism: it is only because the mind has this defence mechanism that projection is possible. The imagination is able to recruit the capacity to project and redeploy it in other contexts; contexts that are not necessarily ones in which a psychological defence mechanism has any obvious role.[10]

(p. 485) Wollheim acknowledges that projection's function as a defence mechanism is lifted from classical psychoanalytical theory. Wollheim (1987) calls this *simple projection*. Simple projection involves the mind projecting a mental state or disposition onto another mind. The consequence of my act of simple projection is that my beliefs about the world have changed (more specifically, my belief about the content of another mind has changed), as have my beliefs about my own mind. For example, in order to alleviate my own feeling of anxiety, I project hostility onto you, and so by imagining it is you rather than I who feels this way, I alleviate my own unpleasant feelings.

Wollheim also develops a concept of *complex projection*, which is central to his account of expressive perception (considered in the following section). He begins with *simple projection*, and supposes that complex projection occurs when simple projection is not possible; that is, when one seeks to project onto an object that one does not conceive of as possessing a mind of its own. In this case, the mind cannot project a mental state or mental disposition, because it knows that only other minds can possess mental states and mental dispositions. So, instead, it projects a *projective property* which is not a mental state or disposition, but which is *of a piece* with the mental state or disposition that the mind would otherwise project. The projection of a projective property onto something that does not possess a mind, Wollheim calls *complex projection*. The essential upshot of complex projection is that the projective property changes how we perceive the object onto which we have projected it, i.e. how the object appears to us is changed. This involves an affective property fusing with a perceptual property, so that sense perception and affect are experienced together. The consequence of my act of complex projection is that the world as it appears to me changes (more specifically, my perception of a portion of nature or a work of art changes by becoming fused with an emotional quality). For example, I am walking along the bank of an estuary and I feel melancholy. I seek to alleviate this feeling through projection, and, as there is no one else to project it onto, I seek to project it onto the estuary itself. I know that the estuary does not possess its own mind, and so I cannot project the melancholy onto the estuary through simple projection. So, instead, I project a projective property that is of a piece with melancholy onto the estuary. My beliefs about the estuary have not changed, but how I perceive it has changed. My perception of the appearance of the estuary, with its slow-moving current

and the weeping branches of the willow trees, fuses with my feeling of melancholy, so that the estuary's appearance seems to have a melancholy quality.

Projection also has a role in Wollheim's understanding of value and morality. Wollheim (1984) distinguishes value from morality, and believes that both can be explained psychologically. He conceives of morality in terms of obligation, and maintains that it has its origin in the introjection of a frightening voice that tells us that we may not act according to our natural impulses—a prospect he finds wholly baneful (Wollheim 1999). Value, on the other hand, has its origin in projection: in order to preserve a state of *archaic bliss* or *love satisfied*, the mind projects this state onto some object (usually a person) in the world. The sense of 'projection' here is given by considering the response to a person upon whom archaic bliss has been projected which involves neither the belief that the person is in a state of bliss, nor the fusing of blissfulness with the perception of the person. (p. 486) Rather, our evaluation of the person changes: we judge the person to possess value on account of the archaic bliss that we have projected onto them.[11]

So Wollheim conceives of different forms or senses of projection: mental states and dispositions can be projected (simple projection); projective properties can be projected (complex projection); and archaic bliss can be projected (love). These different forms of projection have different consequences: our beliefs about other minds change through simple projection; our perception of the appearance of the world changes through complex projection; and our evaluation of other people's value changes through the projection of archaic bliss.

This approach to projection is notable for at least three reasons. First, although Wollheim begins with the Freudian defence mechanism of projection, he develops it in novel ways that exceed anything found in the Freudian or Kleinian corpus. Secondly, it is arguable that Wollheim conceives of projection operating in several different ways: it appears that there is a variety of different forms of projective experience in his philosophy. Finally, projection is recruited in a range of different contexts in a way that draws together his thinking on topics in different branches of philosophy. In doing so, it seems to be central to his very understanding of the human condition.

For Wollheim, human beings are fundamentally creatures who project their inner life onto the world in a range of different ways. No creature that lives in the real world is going to flourish unless the relationship between the creature's mental life and the real world is adequately mediated. Wollheim offers us an insight into the range of ways in which projection mediates our beliefs, evaluations, and perceptions. In doing so, he helps us to get clearer about human flourishing.

On Projection as the Foundation for Expressive Perception

Artistic expression has long been a concern of philosophers who seek to account for what it is that is both distinctive and valuable about art. What makes something count as a work of art has variously been identified with an object's representational properties (Plato 1974), its formal properties (Bell 1913), and its expressive properties (Tolstoy 1898/1995). When it comes to providing an account of the expressive properties of works of art, attempts have been made to link these to the emotion experienced by the artist when creating the work (Tolstoy); or the sense in which a work of art might resemble physiological expression of emotion—just as a Saint Bernard dog's face resembles the look of a sad human face (Kivy 1989); or it might be linked to the work's capacity to arouse an emotion in its audience (Matravers 1998) or to the work's capacity to engage the audience's (p. 487) imagination in a way that produces some kind of emotional response (Walton 1988); or, finally, it might be linked to the idea that artist and audience alike are able to comprehend otherwise incomprehensible emotions through engaging with the work of art (Collingwood 1938).

Such theories of artistic expression are not the only reason that one might seek an account of expressive properties and our capacity for expressive projection. Our ability to perceive emotion in human physiognomy also invites an account of these properties and capacities in the context of self-expression (e.g. Green 2007) and our possible knowledge of other minds (e.g. Hampshire 1971). The other context in which we might encounter expressive properties is in our experience of nature, some portions of which can, from time to time, seem expressive of emotional states—e.g. the estuary that is expressive of melancholy. Whatever it might mean for us to identify a portion of nature with a certain emotion, there is one sense in which this is distinctly different from identifying the quality of a person's countenance, gait, or tone of voice with such an emotion. In the physiognomic case, we are aware that the person possesses a psychology, and that the emotional quality that we perceive in the person's appearance might sensibly be thought to have its genesis in the person's psychology. In the case of nature, such a line of thought is not possible (absent anthropomorphism), because we know that the portion of nature does not possess a psychology to which the emotional quality might be genetically related. This requires us to distinguish between *expression* and *expressiveness* (Budd 2002). An artefact might be an expression of emotion, and this might account for why we perceive it as being expressive of that emotion. But an object might well seem to be expressive of emotion without its being an expression of emotion.

Works of art constitute a special category when it comes to expression and expressiveness. Like physiognomy, artistic creativity gives rise to something that might be an act of self-expression or a 'pressing out' of the artist's inner condition. As with nature, we can engage with a work of art as an object that has properties that affect us independently of anything to do with creative intentions. In §§17–18 of *Art and its Objects*, Wollheim indicates that he is aware of these different possibilities, and

distinguishes between art as an expression and the expressiveness of art. He seems to acknowledge that the emotional qualities can be a pressing out of the inner condition of the artist, but he does not think that this is all that is involved. He maintains that the expressiveness of a work of art has much to do with the audience's capacity to project its own emotions onto the work of art. It is at this point that we find the beginnings of his attempt to use projection to explain our capacity for expressive perception of works of art.[12]

Complex projection provides the basis for his theory of expressive perception. Wollheim believes that we perceive melancholy in the appearance of the estuary because we have projected a projective property that is of a piece with melancholy onto the estuary. (p. 488) He then moves from the expressive perception of nature to the expressive perception of art. Wollheim maintains that we can project projective properties onto any object, and he further maintains that the archaic mind does this when simple projection is not possible. In some cases, the effect of complex projection is short-lived. In other cases, the effect is longer lived, because there is a match between the appearance of the object and the affect that the mind seeks to project. This is because sometimes there is a relationship of matching between what is being projected and that onto which it is projected; which relationship Wollheim calls *correspondence*, drawing on the idea of *correspondance* in Baudelaire's poetry and the visionary Swedenborg's writing.[13]

As a defence mechanism, it does not matter whether the projective property is projected onto an object whose appearance corresponds to it. Relief will be obtained irrespective of this. But correspondence is a requirement for expressive perception. When a work of art's appearance corresponds to the projective property that is projected onto it, the appearance sustains the projective property, such that the appearance of the object invites projection, and will be perceived as being expressive of the affect that was earlier projected onto it. In this way, our perception of the object is changed for as long as the projection is sustained.

It is not enough, however, for Wollheim to explain why we perceive emotional qualities in nature when we have projected our own emotions onto it. The interesting situation is not one in which I feel melancholy, and then perceive the world around me to be melancholy because I have projected my current melancholy onto it. Rather, the interesting case is the one in which I do not currently feel melancholy, but when I encounter a work of art, I become aware that it seems to have a melancholy quality, and this somehow puts me in mind of melancholy. Wollheim (1993) is adamant that correspondence grounded in complex projection can account for this. As a contribution to our understanding of artistic expression, there is something very satisfying about Wollheim's solution. Claims that Constable's *Hadleigh Castle, the Mouth of the Thames—Morning After a Stormy Night* is expressive of harrowing sadness no longer boil down to claims about the audience's knowledge of Constable's harrowing sadness when he painted it, or claims that it somehow resembles the natural expression of a person in the grip of harrowing sadness, or that it simply causes the audience to feel sad in the ordinary sense. Rather, these give way to the claim that the painting's expressiveness is to be explained in terms of a

distinctive mode of perceiving it; such that the sensory perception of the painted surface and the harrowing affect to which it corresponds are fused together in a single experience with a distinctive phenomenology that fuses the sensory and the emotional.

Yet, for all its appeal, we cannot let it escape from the host of criticisms that it meets. First, there are the problems that it presents for classical psychoanalysis. Complex projection is not to be found in the psychoanalytical canon. It appears to be a case of Wollheim extending ideas found in Freudian psychoanalysis in a new and original way. This is an admirable enough project, but it raises worries about whether he has provided (p. 489) enough evidence to justify this extension. (If it were a case of applying standard ideas, it might be sufficient for him to appeal to authority, but this is not so straightforward when he extends those ideas.) So, even if one accepts the legitimacy of psychoanalytical theory, it might be argued that psychoanalysis does not provide philosophical authority for Wollheim's innovation.

Secondly, there are the philosophical objections. Budd (2001) provides a close analysis of Wollheim's theory and its weaknesses. He argues that Wollheim's argument fails to prove that expressive perception has its origin in complex projection; fails to give an adequate account of correspondence; and fails to demonstrate that correspondence is anchored in complex projection. Budd claims that Wollheim relies too much on stipulation, and that he cannot provide an account of complex projection that has the precision required for a philosophical analysis. Wollheim (2001) concedes to Budd that aspects of his theory are imprecise, indefinite, and merely programmatic, but suggests that these are inherent in any theory of expressive perception, at least given the limits of our current understanding of it.

Thirdly, the whole approach might be misguided. Budd, whatever his criticisms, seems to appreciate the appeal of such an account of expressive perception, were Wollheim able to deliver it. It might be objected, however, that it is a mistake to look for such an account. Wollheim maintains that there is something melancholy about the estuary in the sense that something about its appearance actually corresponds to melancholy. This is something more than an imagined relationship between the appearance and the affect. The Pathetic Fallacy espoused by Ruskin (1856), however, denies that there is anything other than an imaginary connection between melancholy and the estuary, and so a theory, such as Wollheim's, that attempts to capture how it can be that we experience the appearance of nature as imbued with some emotional property is misguided.

Finally, there is another line of objection that perhaps trivializes the extent of Wollheim's achievement. The account of expressive perception is supposed to demonstrate that perception and affect can be fused together, and the expressiveness of art is a special case in which this occurs. Arguing that such a fusion is possible, and that it is a special feature of the experience of art seems to be innovative. Collingwood (1938: 162) observes, however, that, at least in the case of colour, the fusing of perception and emotion is quite natural, and so it only seems remarkable to 'adult and "educated" people

in what is called modern European civilization', who are taught to sterilize their colour perception of any emotional charge.

On Psychoanalysis as an Aid to Understanding Art

Aside from questions about the nature of art and the kinds of properties that works of art have, a central problem of the philosophy of art concerns our understanding of the meaning of a work of art. This is the problem of art criticism. We can ask questions about (p. 490) what artists do when they create works of art, and what audiences do when they engage with works of art. There is also a special category of spectator that is the critic. The critic professes to have a special appreciation of the meaning of works of art, and the ability to communicate this to others, so that they, too, might come to appreciate the value that the critic finds in the work of art.

The traditional position maintains that art criticism is a matter of understanding the meaning of a work of art. Every work of art has a meaning, and this is so because every work of art is an intentional object: it was created by an artist who had some specific intention as to the meaning that the object would have. So there are right and wrong interpretations of the meaning of a work of art. There is a standard of correctness that determines whether an interpretation is correct or not, and that standard is the artist's intention: only interpretations that are consistent with the artist's intention are correct. So the challenge for the critic is to reveal the artist's intention in creating a specific work of art.

This understanding of criticism was challenged in the twentieth century by the idea that the true meaning of the work of art is the meaning that it in fact has for any given audience; not the meaning that the artist intended it to have (Richards 1925). On this understanding of criticism, the critic's role is not to retrieve the artist's intention, but to articulate the meaning that the work has for the critic. Different people will experience a work differently, and that is fine. It is only the properties of the work of art that determine whether interpretations are acceptable, and the critic's task is to link the subjective interpretation to the objective properties of the work of art. The imperative to retrieve the artist's intention in order to determine the correct interpretation of the work is now gone.

Wollheim rejects this innovation. He believes that there is a standard of correctness for understanding the meaning of a work of art, and it lies with the artist. He does not maintain, however, that it is a matter of retrieving the artist's intention. In 'Criticism as Retrieval', Wollheim (1980) explains that the critic's role is to retrieve the creative process that terminates on the work of art—not the artistic intention behind the work of art. This, he maintains, determines the correct meaning of a work of art. The creative process is more expansive than the artist's intentions, because it includes the vicissitudes to which the artist's intentions are subject. Such vicissitudes may be intentional (such as

changes of intention), but others involve chance or elements that are beyond the artist's control. The creative process is also said to include background beliefs, conventions, and modes of production against which the artistic intention is formed. Wollheim maintains that theory and hindsight may be used to reconstruct the creative process, and finds archaeology to be a useful metaphor for criticism as retrieval.

Wollheim engages in criticism as retrieval in *Painting as an Art*. When he engages with works by the masters, he takes it for granted that psychological factors deep within the artist's psyche are often at work in a painting, and that it is possible to retrieve these when attending carefully to paintings. Wollheim believes that any painter who is genuinely an artist paints with the intention of imbuing his canvas with pictorial meaning. For Wollheim, pictorial meaning is concerned with the artist's attempt to explore the (p. 491) depths of his emotional life, and, in doing so, he argues for the 'repsychologization' of pictorial meaning. He provides compelling readings of the *Antichus and Stratonice* paintings, to make the case that Ingre paints out of an unconscious wish to reorganize his family drama; that Poussin explores morality in paintings such as *Rinaldo and Armida* and *Armida Carrying off Rinaldo*; and that De Kooning is concerned with infantile experiences of pleasure that shape subsequent knowledge and desire in such paintings as *Untitled III, 1977, Untitled II, 1979*, and *Woman I*.

All of this clearly draws on psychoanalysis. Yet Wollheim is insistent that he does not engage in psychoanalysis in the book (1987: 8–9):

> To the experience, to the hard-won experience, of painting, I then recruited the findings of psychology, and in particular the hypotheses of psychoanalysis, in order to grasp the intention of the artist as the picture revealed it. In these lectures I have not attempted to 'psychoanalyze' individual painters, but, if I am right in thinking that art presupposes a common human nature, and that pictorial meaning works through it, then it must be absurd to bring to the understanding of art a conception of human nature less rich than what is required elsewhere. Many art-historians, in their scholarly work, make do with a psychology that, if they tried to live their lives by it, would leave them at the end of an ordinary day without lovers, friends, or any insight into how this had come about.

Wollheim may not actually psychoanalyse artists, however he comes fairly close to doing so. He certainly believes that the meaning of great art lies in the attempts of great artists to engage with the struggles at the deepest parts of their psyche. Whether or not he succeeds in showing that an artist can explore such concerns through the painted surface, it is far from certain that these are the only legitimate concerns that an artist might seek to explore: a great artist might, for instance, be interested in social commentary or experimenting with the artistic medium without having an interest in self-analysis.[14]

On Wollheim as a Guide to Art, Philosophy, Psychoanalysis, and Life

Why does Wollheim matter for understanding the relationship between philosophy, psychoanalysis, and art?

First, he demonstrates the dynamic nature of this relationship. For Wollheim, art is not merely an object of aesthetic appreciation, and certainly not merely an object of scholarly rumination. It was integral to the way he lived his life. One cannot neatly separate Wollheim's creative writing, his criticism, or his theoretical writing from one (p. 492) another or from his life. That, in itself, is telling for how philosophers of art might approach their subject more fruitfully. Psychoanalysis was central to his life, writing, and appreciation of art, and to the way he made sense of these. Indeed, one cannot separate out Wollheim's analysis with Dr Sohn from his theoretical writing about psychoanalysis, or from the way he philosophized, or lived his life. Psychoanalysts will not find this remarkable, because it is a feature of their training that they must undergo an analysis in order to analyse others, and on account of the fluidity of the relationship between therapy and theory in psychoanalysis. It is a more remarkable thought, however, for the analytical philosopher, and one worth dwelling on.

Secondly, Wollheim exemplifies the riches of pursuing a psychoanalytically informed philosophy, as opposed to philosophy of psychoanalysis. One of the unifying themes that we can see running throughout his psychoanalytically informed philosophy is projection as an activity that is central to being human. This comes out of his engagement with psychoanalysis, but it moves beyond classical psychoanalysis, and is seen in his treatment of artistic expression, which turns out to depend upon our capacity for expressive perception, which is anchored in a modification of the classical psychoanalytical defence mechanism of projection.

Finally, he restores the primacy of the artist to the meaning of a work of art, but, in doing so, he adds a psychoanalytical twist: meaning is no longer the artist's intention, but the process that he undergoes in creating the work of art. At least in the case of painting, the pictorial meaning that is retrieved in this process is psychological, and Wollheim demonstrates how the hypotheses of psychoanalysis may be used to understand the deep psychosexual issues being explored by the artists through their manipulation of the painted surface.

In each of these respects, Wollheim provides thoughtful contributions that challenge established approaches to the subject. That is not to say that he has had a decisive impact on current thinking. The history of twentieth-century philosophy may well treat him as little more than a curiosity, a footnote in the development of the major lines of thought. His writing offers a most valuable inspiration, however, for readers who are captured by the intensity of certain experiences, and the significance that these experiences have for us in virtue of their intensity. It demonstrates how one might think and write about such experiences in a rigorous and analytical way—a challenge that is notoriously difficult

when it comes to discussing the phenomenology of experience. Wollheim did this, in part, through his training as an analytical philosopher. In major part, it was also because of his ability to recruit the insights of psychoanalysis in a liberal way that created a psychoanalytically informed philosophy.[15]

References

Ayer, A. J. (1962). *Language, Truth and Logic*. New York: Dover.

Baumgarten, A. G. (1750). *Aesthetica*. Hildesheim: Georg Olms.

Bell, C. (1913). *Art*. London: Chatto & Windus.

Blackburn, S. (1985). *Spreading the Word*. Oxford: Oxford University Press.

Budd, M. (2001). 'Wollheim on correspondence, projective properties, and expressive perception'. In R. van Gerwen (ed.), *Richard Wollheim on the Art of Painting* (pp. 101–111). Cambridge: Cambridge University Press.

Budd, M. (2002). *The Aesthetic Appreciation of Nature*. Oxford: Oxford University Press.

Budd, M. (2005). 'Richard Arthur Wollheim: 1923–2003'. *Proceedings of the British Academy 130*: 227–246.

Bullough, E. (1912). ' "Psychological distance" as a factor in art and an aesthetic principle'. *British Journal of Psychology 5*: 87–98.

Collingwood, R. G. (1938). *The Principles of Art*. Oxford: Oxford University Press.

Danto, A. (1964). 'The artworld'. *Journal of Philosophy 16*: 571–584.

Dickie, G. (1974). *Art and Aesthetic*. Ithaca, NY: Cornell University Press.

Freeman, D. (2012). *Art's Emotions: Ethics, Expression and Aesthetic Experience*. Durham: Acumen.

Freeman, D. (2016). 'Human complexity and *The Thread of Life*'. In J. Malpas and A. Wojtowicz (eds), *On Human Complexity, Proceedings of the 2015 W. D. Joske Interdisciplinary Colloquium*, Hobart. Sandy Bay, Tasmania: University of Tasmania School of Humanities.

Freud, S. (1895). 'Project for a scientific psychology'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 1 (pp. 283–397). London: Hogarth Press.

Freud, S. (1911). *Psycho-Analytic Notes on an Autobiographical Account of a Case of Paranoia*. In J. Stratchey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 12. London: Hogarth Press.

Gerwen, R. van (ed.) (2001). *Richard Wollheim on the Art of Painting*. Cambridge: Cambridge University Press.

Green, M. S. (2007). *Self-Expression*. Oxford: Clarendon Press.

Hampshire, S. (1971). *Freedom of Mind*. Princeton, NJ: Princeton University Press.

Hegel, G. W. F. (1835/1975). *Aesthetics. Lectures on Fine Art*, 2 vols, trans. T. M. Knox. Oxford: Clarendon Press.

Hepburn, R. (1966). 'Contemporary aesthetics and the neglect of natural beauty'. In B. Williams and A. Montefiore (eds), *British Analytical Philosophy* (pp. 285–310). London: Routledge.

Herwitz, D. (1991). 'The work of art as psychoanalytical object: Wollheim on Manet'. *Journal of Aesthetics and Art Criticism 42*: 137–153.

Hopkins, J. and Savile, A. (eds) (1992). *Psychoanalysis, Mind and Art*. Oxford: Blackwell.

Hume, D. (1739–40/2000). *A Treatise of Human Nature*, eds D. F. Norton and M. J. Norton. Oxford: Oxford University Press.

Kant, I. (1790/1987). *The Critique of Judgment*, trans. W. S. Pluhar. Indianopolis, IN: Hackett Publishing.

Kivy, P. (1989). *Sound Sentiment*. Philadelphia, PA: Temple University Press.

Lear, J. (2005). 'The bodily ego: memorial to Richard Wollheim'. Presented to American Philosophical Association (Pacific Division), 27 March 2005 (unpublished).

Lear, J. (2006). *Radical Hope*. Cambridge, MA: Harvard University Press.

Marcus, L. (2015). 'Autobiography and psychoanalysis'. In Z. Leader (ed.), *On Life-Writing* (pp. 257–283). Oxford: Oxford University Press.

Matravers, D. (1998). *Art and Emotion*. Oxford: Clarendon Press.

Plato (1974). *Republic*, trans. G. M. A. Grube. Indianapolis, IN: Hackett.

Podro, M. (2004). 'On Richard Wollheim'. *British Journal of Aesthetics 44*: 213–225.

Price, A. W. (1992). 'Three types of projectivism'. In J. Hopkins and A. Savile (eds), *Psychoanalysis, Mind and Art* (pp. 110–128). Oxford: Blackwell.

Richards, I. A. (1925). *The Principles of Literary Criticism*. London: Kegan Paul, Trench, and Trubner.

Ruskin, J. (1856). 'Of the pathetic fallacy'. In *Modern Painters*, Vol. 3, Part 4. London: Smith, Elder & Co.

Shaftsbury, A., Earl of (1900). *Characteristics*, ed. Robinson. London: G. Richards.

Stokes, A. (1972). *The Image in Form: Selected Writings of Adrian Stokes*, ed. R. Wollheim. New York: Icon Editions.

Tolstoy, L. (1898/1995). *What is Art?*, trans. R. Pevear and L. Volokhonsky. Harmondsworth: Penguin.

Velleman, J. D. (2005). *Self to Self*. Cambridge: Cambridge University Press.

Walton, K. (1988). 'What is abstract about the art of music'? *Journal of Aesthetics and Art Criticism 46*: 251–264.

Wollheim, R. (1959). *F. H. Bradley*. Harmondsworth: Penguin.

Wollheim, R. (1961). *Socialism and Culture*. London: Fabian Society.

Wollheim, R. (1964). 'A paradox in the theory of democracy'. In P. Laslett and W. G. Runciman (eds), *Philosophy, Politics, and Society* (pp. 71–87), Second Series. Oxford: Basil Blackwell.

Wollheim, R. (1969). *A Family Romance*. London: Jonathan Cape.

Wollheim, R. (1972). 'Introduction'. In *The Image in Form: Selected Writings of Adrian Stokes* (pp. 9–31). New York: Icon Editions.

Wollheim, R. (1974). *On Art and the Mind*. Cambridge, MA: Harvard University Press.

Wollheim, R. (1980). *Art and its Objects: With Six Supplementary Essays*, 2nd edn. Cambridge: Cambridge University Press.

Wollheim, R. (1984). *The Thread of Life*. New Haven, CT: Yale University Press.

Wollheim, R. (1987). *Painting as an Art*. London: Thames and Hudson.

Wollheim, R. (1991). *Freud*, 2nd edn. London: Fontana Press.

Wollheim, R. (1993). *The Mind and Its Depths*. Cambridge, MA: Harvard University Press.

Wollheim, R. (1999). *On the Emotions*. New Haven, CT: Yale University Press.

Wollheim, R. (2001). 'A reply to the contributors'. In R. van Gerwen (ed.), *Richard Wollheim on the Art of Painting* (pp. 241–262). Cambridge: Cambridge University Press.

Wollheim, R. (2004). *Germs*. Church Enstone: Waywiser Press.

Notes:

(^{1}) For an analytical aesthetician's perspective on Wollheim's life and work, see Budd (2005); for a psychoanalyst's perspective, see Lear (2005); and for an art historian's, see Podro (2004).

(^{2}) For further discussion, see Marcus (2015).

(^{3}) For a survey of responses to Wollheim's philosophy of painting, see the various contributions to *Richard Wollheim on the Art of Painting* (Gerwen 2001).

(^{4}) For further discussion, see Freeman (2016).

(^{5}) His knowledge of psychoanalysis was recognized by the British Psychoanalytical Society (which made him an honorary affiliate in 1982), the San Francisco Psychoanalytical Institute (which made him an honorary member in 1994), and the International Society for Psychoanalysis (which awarded him for his distinguished services to psychoanalysis in 1991)—rare honours for those who have not undergone formal psychoanalytical training.

(^{6}) It is in keeping with his earlier *F. H. Bradley* (1957), in which he provides an exposition of the British idealist's metaphysical system and subsequent misunderstandings of it. For a comparison between aspects of British idealism and psychoanalysis, see Wollheim's Dawes Hicks Lecture on Philosophy at the British Academy, 'The Good Self and the Bad Self: The Moral Psychology of British Idealism and the English School of Psychoanalysis Compared' in *The Mind and Its Depths*. For a discussion of F. H. Bradley and T. S. Eliot, see Wollheim's 'Eliot and F. H. Bradley' in Martin, G. (ed.), *Eliot in Perspective* (London, 1972); reproduced in *On Art and the Mind*.

(^{7}) Wollheim's most significant contribution to political philosophy was 'A Paradox in the Theory of Democracy' (Wollheim 1964). In the realm of social and political thought, he produced *Socialism and Culture* (Wollheim 1961), a pamphlet for the Fabian Society. In terms of the relationship between politics, art, and psychoanalysis, his most fascinating line of thought is the enigmatic last paragraph of *Painting as an Art* (1987: 357), when he concludes his argument for the psychological meaning of painting thus: 'the societies across which painting has survived as an art are human societies: that is, societies in which a common human nature manifests itself … Our belief then in the value of painting … happily locks together two of the commitments by which I steer: the love of painting, and loyalty to socialism'.

(^{8}) For a range of perspectives on Wollheim's philosophy, see the various contributions to Hopkins and Saville (1992).

(^{9}) Projection comes into philosophy in an important way as a means of understanding the distinction between fact and value. For Hume, the good and the beautiful are not part of how the world actually is. They are not properties of objects, but rather are projected onto objects by subjects. For a subject to claim that an object is beautiful is not to assert

that it possesses some objective property, but that the subject feels a certain way about the object. Hume speaks of the way that we *stain* the world with value. So claims about value are not claims about facts. In the twentieth century, philosophy famously took the 'linguistic turn', and its focus became linguistic analysis. A. J. Ayer (1962) established a way of analysing how we use language when we make value claims, and this classical statement of logical positivism is subsequently developed by Simon Blackburn (1985). In some sense, this is simply a development of Hume's insight. But now the focus has shifted from understanding the relationship between fact and value (a metaphysical problem), to understanding the linguistic structure of value claims.

([10]) For a comparative discussion of projection in the philosophy of Blackburn, J. L. Mackie, and Wollheim, see Price (1992).

([11]) For a discussion of morality, obligation, and love, see Backstrom (this volume), who presents an alternative to Wollheim's understanding here.

([12]) The following is a sketch of how Wollheim grounds artistic expression in a psychoanalytical notion of projection, but for a detailed discussion of Wollheim's approach to the expression of emotion in art, see Freeman (2012). Wollheim's earliest essay on the topic is 'Expression' (in Wollheim 1974) and the mature position is found in Wollheim (1987) and 'Correspondence, Projective Properties, and Expression in the Arts' (in Wollheim 1993).

([13]) A discussion of correspondence is crucial to understanding Wollheim's theory of expressive perception, but is beyond the scope of the present chapter. See Wollheim (1999: 78–80, 1987: 81–4, and 1993: 154–5), as well as Freeman (2012: 59–61) for a discussion of this concept.

([14]) For a critical evaluation, see Herwitz (1991).

([15]) Beyond aesthetics, his legacy might be seen in the inspiration that he gave a subsequent generation of American philosophers who have adopted a psychoanalytically informed approach to philosophy, such as may be witnessed in the treatment of philosophical anthropology and ethics, in Jonathan Lear's *Radical Hope* (Lear 2006), and the philosophy of mind and psychology in some of the essays in J. David Velleman's *Self to Self* (Velleman 2005).

Damien Freeman

PM Glynn Institute, Australian Catholic University, Australia

Oxford Handbooks Online

Literary Form and Mentalization ⓐ

Elisa Galgut

The Oxford Handbook of Philosophy and Psychoanalysis
Edited by Richard G.T. Gipps and Michael Lacewing

Abstract and Keywords

This chapter argues that literature—or at least certain kinds of literature—facilitates mentalization. Book reading shares some of the same features as mind reading. Psychoanalysis, via the work of Hanna Segal and others, has been interested in the differences between escapist literature and literature that encourages genuine psychological engagement. This chapter engages that issue via the lens of mentalization. It focuses specifically on literary form rather than on content, and examines the ways in which some kinds of literary form facilitate mentalizing capacities. More narrowly, it shows how different kinds of literary techniques—the free indirect discourse employed by Jane Austen, the tight structure of the sonnet form—enable different mentalizing abilities and develop our capacity for self-reflection.

Keywords: psychoanalysis, mentalization, self-reflection, literature, literary form, free indirect discourse, sonnet, Jane Austen

Elisa Galgut

(p. 495) Introduction

SINCE its beginnings, psychoanalysis has been used to explore our engagement with art in general and literature in particular; indeed, the explorations have gone both ways, with literature used as a resource for psychoanalytic insight. Early psychoanalytic literary analysis tended to focus on the *content* of literary narratives—the fate of Oedipus, the indecision of Hamlet, the typical plots of romance tales where 'all the women in the novel invariably fall in love with the hero' (Freud 1908: 150). Freud does nod towards 'the purely formal—that is, aesthetic—yield of pleasure' (1908: 153) but he provides little analysis of these formal elements. Hanna Segal, in her seminal 1952 paper, takes a significant step in shifting the discussion away from the emphasis on content by exploring

its relationship with literary *form*. Segal develops an account of the *reparative* function that aesthetic formal elements can play in tragedy in particular, and art more broadly. The orientation of this chapter follows Segal in falling broadly within the British object-relations tradition, which I largely take for granted.[1] But I take the conversation between psychoanalysis and literature in a slightly different direction. Much psychoanalytically informed art criticism focuses primarily on the relationship between artworks and *unconscious* mental processes—the ways in which art gives expression to unconscious fantasy, for example, or its relation to dreaming[2] or daydreaming. I do not dispute the importance of the unconscious in both the production and appreciation of art. However, my focus in this chapter explores more closely the ways in which formal elements of literature can facilitate higher-order mental functioning.

(p. 496) The philosopher Jenefer Robinson has argued that 'formal or structural devices in literature play the role of coping mechanisms' by directing 'the sequence of appraisals and reappraisals that we engage in as we read, and in particular, they act as coping strategies … which enable us to manage or deal with the explosive fantasies' (Robinson 2005: 196) found in the content of the work. In this chapter, I too focus on the importance of literary form by showing some of the ways in which formal properties enable our capacities for mentalization. Although it is likely the case that even basic engagement with narrative stories facilitates the capacity for mentalization,[3] I shall argue that particular forms of literary techniques are better suited for the facilitation of mentalizing capacities, particularly the capacity for meta-cognitive functions. This is so because literature provides us with the opportunities for insight and focused attention not often afforded to us in ordinary life. Here I focus on two forms of literary technique: the stylistic feature known commonly as 'free indirect discourse' and the poetic form of the sonnet. This discussion is not intended to be exhaustive of the ways in which literary form may facilitate mentalization, but rather to gesture in the direction in which new insights from psychoanalytic thought shed light on how reading literature—and especially good literature—enables the development of cognitive and emotional capacities. There has been a surge in the analysis of readers' responses to fiction using the insights of cognitive science, with particular reference to simulation theory to explain how we engage with the minds of (fictional) others.[4] I would like to make a case for expanding our accounts of engaging with literary texts via insights gleaned from psychoanalysis. By doing so I am also making a claim for a humanistic view of literature, which holds that literature can tell us important things about ourselves as thinking and reflective subjects.

A Word about Mentalization

Mentalization—and its related verb *mentalizing*—was reintroduced into the contemporary psychoanalytic literature by Fonagy, Target et al. in their groundbreaking book on mentalization and its relation to development. 'Mentalization' refers to our ability to understand those around us and to predict their behaviour, as well as the ability to reflect upon our *own* minds and actions. There is some discussion in the literature about whether the term adds something new to psychology—and particularly to psychoanalysis—or

whether it's just old wine in new bottles. Cognitive psychologists, for instance, have been using terms such as 'theory of mind' and 'simulation' to describe similar skills. I argue that the terms *mentalization* and *mentalizing* add a dimension to our theoretical understanding of how we understand minds by focusing on certain meta-cognitive abilities which are both amenable to and enriched by the incorporation of psychoanalytic theory. (p. 497) Mentalization as a theory of mind also grounds mentalizing capacities firmly within affective capacities, which cognitive science approaches tend not to do.

The psychoanalyst Jeremy Holmes has usefully delineated four interrelated aspects of mentalization:

> **1.** Mentalizing is a 'meta-cognitive' phenomenon—*thinking about thinking*—and thus requires higher-order beliefs.
> **2.** 'Mentalizing is concerned with the *meanings* which we attribute to our own and others' actions—that is, to the implicit or explicit hypotheses we use to understand why we, or another, might have thought or done such and such a thing.'
> **3.** Mentalization is a key aspect of *persons*—one might say that it is both central to how we respond to another *as* a person, as well as an important constituent of *personhood*.
> **4.** 'Mentalizing is not a fixed property of mind, but is a *process*, a capacity, or skill, which may be present or absent to a greater or lesser degree' (Holmes 2005: 180).

There has been some discussion about whether the notion of mentalization is sufficiently psychoanalytic: does it differ from—say—appealing to a theory of mind in order to explain how we engage with and understand others? Mary Target (2008) addresses this issue, and argues that mentalization *is* part of psychoanalysis 'understood in a broad sense'.

> Mentalization as a process means thinking about what someone does, and what happens between people ... in terms of psychological meaning and motivation. Psychoanalysis clearly contributes unique understanding of particular kinds of motivation ... The theory of stages of mentalization underlying the development of personality and relatedness ... is psychoanalytic in that it concerns the unconscious foundations of mental functioning, which may be dynamically as well as descriptively unconscious.

> (Target 2008: 262)

I am in sympathy with Target's understanding of mentalization as falling within a psychoanalytic understanding of the mind; indeed, part of the motivation for its adoption is to distinguish it from other kinds of terms (such as 'theory of mind') regarding how we understand the minds of others. Mentalization, in both its theoretical origins and its clinical use—its relationship with attachment theory, its use in the treatment of patients diagnosed with borderline personality disorder—is steeped in ways of understanding the mind that are psychoanalytic at heart. This indicates a fundamental difference between the concept of mentalization and that of theory of mind as used within cognitive science. The two concepts are clearly *compatible*, but discussions about theory of mind in

cognitive science tend to refer mainly to mental states involving propositional attitudes—beliefs and desires—without locating this discussion in a particular psychological theory. Theory of mind 'fails to encapsulate the relational and affect regulative aspects of interpreting behaviour in mental state terms' (Fonagy 2008: 3). As noted earlier, mentalization is not only a meta-cognitive phenomenon, but it is also a meta-emotional (p. 498) one—it refers to an ability not only to think *about* emotion, but to use emotion to think about thinking.[5] Mentalization thus points to the ways in which emotion and cognition are interdependent—'comprehension of intentional mental states encompasses the understanding of feelings, and the representations which underlie them' (Diamond and Kernberg 2008: 236). Elliot Jurist claims that affectivity—the capacity to reflect on affective and emotional states—is 'a specific aspect of mentalization' (Jurist 2005: 428); Jurist stresses the importance to mentalization of such factors as the attachment bond between infant and caregiver, and the relationship between the development of affect and the development of a sense of self. Jurist concludes that mentalization 'has an intrinsic relation to affects' and 'fosters a new, more differentiated kind of affect and self-regulation' (2005: 428). In the discussion that follows, I hope to show how certain kinds of literary form foster mentalization by requiring the reader to develop not only meta-cognitive capacities, but meta-emotional ones as well.

Another way in which the concept of mentalization differs from a cognitive psychology picture of the mind is that both theory of mind and its rival, offline simulation theory, rarely refer to unconscious mental phenomena, and offline simulation theory often makes use of sub-personal explanations to explain our ability to understand and predict the behaviour of others. Mentalization, on the other hand, is a *personal* level ability—as Holmes notes, the term is used to refer to the ways in which we attribute meaning to ourselves and to others. The implication that follows from this is that mentalization flows from a view of the self as a moral and emotional subject.[6] This personal level view of the mind is consistent with the claim that cognitive and emotional capacities are interrelated—mentalization may refer to different aspects of the mind, such as thinking or imagining, for example, but a personal level approach sees these different aspects as interrelated aspects of a whole self. This is consistent with a psychoanalytic understanding of the mind—as Target says, broadly understood.

By understanding the concept of mentalization psychoanalytically, as opposed to through the lens of cognitive science, we understand in more nuanced ways what may impede or foster it. For instance, a psychoanalytically informed concept of mentalization can explore the ways in which overly rigid defence mechanisms may hinder mentalization. Or an appeal to Freud's distinction between primary and secondary processing can explain how primary modes of thinking dominated by unconscious phantasy and irrational defensive strategies may overwhelm our capacities to mentalize, while secondary processing governed by rationality and responsiveness to reality may enhance them. Mentalization requires not just an unconscious response to another's state of mind but an ability to think about such responses. Mentalization is indicative of one's ability to think about

one's own inner life as well as the inner lives of others; it indicates a capacity for flexibility in judgement and an ability to distinguish between fantasy and reality.

(p. 499) Mentalization and Literature: Narrative Voice

Mentalization provides a useful lens through which to examine the ways in which literature affects us cognitively and emotionally. Although she does not use the term 'mentalization', the philosopher Susan Feagin, in her account of the pleasure we take in tragedies, argues for the importance of meta-cognitive responses. Feagin's explanation of our experience of tragic pleasure is to argue that our enjoyment of the *aesthetic* representation of events involves a meta-response that is—and should be—absent from our engagement with such events were they to happen in real life. Tragic pleasure is thus:

> a meta-response, arising from our awareness of, and in response to, the fact that we do have unpleasant direct responses to unpleasant events as they occur in the performing and literary arts. We find ourselves to be the kind of people who respond negatively to villainy, treachery, and injustice. This discovery, or reminder, is something which, quite justly, yields satisfaction.
>
> (Feagin 1983: 98)

Although Feagin proposes that this meta-response is partly constitutive of our pleasure in tragedy, it is my contention that literature more generally can enable the development of these kinds of meta-responses, which constitute an important aspect of mentalization. I hope to show how literature—via a close examination of literary form and technique—can enable better mentalization by heightening our awareness of the development of our thought processes. Different literary forms or techniques may enable these capacities in different ways. In the following discussion, I explore two of the ways in which this might occur.

There are a variety of means by which novelists can make a reader aware of the mental states of characters: she can, *inter alia*, adopt the position of omniscient narrator; she can present the thoughts of a character in the first person, or she can quote what a character says. All of these techniques require some kind of mentalizing capacities on the part of the reader. But some, I argue, are better than others. Compare these three passages:

> **A:** He was unspeakably happy with her; she governed his household so cleverly and economically that they seemed to live in luxury. She was full of attentions for her husband, spoiling and coaxing him, and the charm of her person was so great that six years after their marriage M. Lantin discovered that he loved his wife even more than during the first days when he knew her.
>
> (Guy de Maupassant, 'The Jewels' (1956): 370)

B: Since he belonged, even at the age of six, to that great clan which cannot keep this feeling separate from that, but must let future prospects, with their joys and sorrows, cloud what is actually at hand, since to such people even in earliest childhood any turn in the wheel of sensation has the power to crystallise and transfix the moment (p. 500) upon which its gloom or radiance rests, James Ramsay, sitting on the floor cutting out pictures from the illustrated catalogue of the Army and Navy Stores, endowed the picture of a refrigerator, as his mother spoke, with heavenly bliss.

(Virginia Woolf, To the Lighthouse (1996): 9)

C: As Mr Shepherd perceived that this connexion of the Crofts did them no service with Sir Walter, he mentioned it no more; returning, with all his zeal, to dwell on the circumstances more indisputably in their favour; their age, and number, and fortune; the high idea they had formed of Kellynch Hall, and extreme solicitude for the advantage of renting it … It succeeded, however; and though Sir Walter must ever look with an evil eye on anyone intending to inhabit that house, and think them infinitely too well off in being permitted to rent it on the highest terms, he was talked into allowing Mr Shepherd to proceed in the treaty, and authorising him to wait on Admiral Croft … Sir Walter was not very wise; but still he had experience enough of the world to feel, that a more unobjectionable tenant, in all essentials, than Admiral Croft bid fair to be, could hardly offer.

(Jane Austen, Persuasion (1997): 17)

The de Maupassant short story (passage A) is told from the perspective of the omniscient narrator; we are privy to the thoughts and feelings of M. Lantin from this external perspective (and later in the story via dialogue). We are also given the reasons for M. Lantin's happiness—his wife's cleverness, attention, spoiling, and coaxing—but we rely on the viewpoint of the narrator for our understanding of M. Lantin. De Maupassant, according to the literary critic Wayne Booth, 'prided himself on writing "objectively" ' (Booth 1983: 184), and this objectivity is apparent in the extract above. Understanding M. Lantin's happiness does require some mentalizing capacities, but not very sophisticated ones—the reader is *told* what he is feeling, and why he is feeling it. Comparing the de Maupassant extract with that from Woolf's *To the Lighthouse* (passage B) reveals a difference in the mentalizing capacities that are required. In the Woolf passage, the reader is asked to reflect more carefully on the mental states of six-year-old James Ramsay. This is done via the lens of the omniscient narrator, whose voice presents a subtlety and complexity to what, in the hands of a more prosaic writer, would have been a simple description of a child's excitement about an anticipated outing. The reader is provided with insight into the complex ways in which mental states may influence one another, but this insight is the author's, not the child's. We are removed, momentarily, from the scene as we inhabit the author's bird's eye view, and then are returned to it. The reader's understanding of young James is seen through the lens of the author's views about human psychology, as evidenced by phrases such as: 'Since he belonged … to that

great clan' and 'since to such people even in earliest childhood'. In this passage Woolf is presenting the psychology of the young James Ramsay through a theoretical lens of folk psychology; we are made privy to his mental states, but at a distance—we don't really see the world through his eyes. The extract from Woolf's novel is illustrative of writing that requires a more advanced form of mentalizing than required by the de Maupassant extract, but it does so from the point of view of an omniscient narrator whose perspective remains outside that of her character's.

(p. 501) The Austen piece (passage C) provides examples of what's known as 'free indirect discourse' or 'narrated monologue'; this is the literary technique whereby the first-person thoughts of a character are written in the grammatical third person. Through free indirect discourse, 'the immediacy of the inner voice' of a character is relayed to us by the narrator who is not also a character in the novel (Cohn 1966: 104). Free indirect discourse, especially in the hands of a great writer, allows for the presentation of a multiplicity of perspectives. It works most powerfully when invisible and provides the author with stylistic flexibility. As the literary critic James Wood notes: 'Thanks to free indirect style, we see things through the character's eyes and language but also through the author's eyes and language, too. We inhabit omniscience and partiality at once' (Wood 2008: 11). In the Jane Austen passage, we see how this habitation of both omniscience and partiality provides the reader with a multiplicity of perspectives on the minds of the characters and by doing so facilitates mentalization.

Jane Austen's *Persuasion* narrates the reunion of Anne Elliot and Captain Wentworth. Years earlier, before the novel begins, Anne was persuaded by a friend of the family, Lady Russell, to reject Wentworth's early proposal of marriage on the grounds that the match was not worthy of her. Lady Russell acts as a mother figure to Anne, whose own mother died when Anne was a teenager. Despite being in love with Wentworth, Anne agrees, and Wentworth departs for the navy, angry and heartbroken. Eight years later Wentworth returns, and *Persuasion* narrates their reunion as they chart their way to a second chance of happiness. The context for Passage C is as follows: Anne's father, Sir Walter, part of the landed gentry but fallen on hard times, is considering renting his family home, Kellynch Hall, to Admiral Croft. Admiral Croft, a naval officer and thus held in low esteem by Sir Walter, is married to Wentworth's sister. Mr Shepherd, Sir Walter's lawyer, is trying to persuade Sir Walter to go through with the lease arrangement.

Passage C is written in the third person, but the views are not only the narrator's—they include those of other characters in the novel. Mr Shepherd, aware that Sir Walter is unhappy about renting out his family home, especially to a naval officer related to the rejected Captain Wentworth, is eager to list the Crofts' merits as tenants. It is he—and not the narrator—who considers the Crofts' 'age, and number, and fortune' to be 'indisputably in their favour', and it is he who claims that they had formed a 'high idea' of Kellynch Hall. Through the author's voice we hear the shrewd sales pitch of the lawyer. This provides us not only with Mr Shepherd's own views, but also with the narrator's assessment of him as a wily lawyer. The narrative then shifts to the voice of the narrator ('It succeeded, however'), and then presents Sir Walter's perspective ('Sir Walter must

ever look with an evil eye'). The narrator's omniscient perspective returns with the words 'Sir Walter was not very wise', which is quickly followed by a qualification ('but still he had experience enough of the world'), which again shows Sir Walter's own perspective—we see how Sir Walter is looking for reasons to accept the Crofts as tenants because he needs the income, despite his reservations about their station. This is also reflected in the convoluted language of the final sentence of passage C, which shows Sir Walter's internal wrestle between pride and need. In this short passage we are presented (p. 502) with several viewpoints: Mr Shepherd's views of Sir Walter and Admiral Croft; Sir Walter's views about himself and Admiral Croft; and the narrator's views. We are made privy to Sir Walter's anxieties about social status, which the narrator mocks, although it is a mockery tempered with sympathy. What is even more remarkable about this particular passage is that these are minor characters, and yet the narrative voice subtly shifts perspectives and encourages reflection on all their points of view,[7] even those with whom we don't sympathize.

Austen's use of free indirect discourse or narrated monologue is subtle, and the reader must pay careful attention to the changes in tone and voice. The reader is asked not only to think about the respective viewpoints as they are presented to us, but is also asked to draw implicit conclusions regarding how the narrator—who, in Austen's world, functions as a moral ballast—views her characters. Austen's use of free indirect discourse allows the narrative flow to remain uninterrupted as it directs the reader's gaze without drawing attention to itself. 'In *Persuasion*, the narrative perspective shifts more often and more quickly than in *Emma*, and the multiple perspectives serve to obscure the origin of judgement. Multiple perspectives are presented, so in one scene the perspectives of all characters involved are simultaneously presented' (Chen 2014: 37).

In another passage, Jane Austen uses free indirect discourse at a significant point in the novel when Anne realizes that Wentworth may still have feelings for her. Wentworth, noting that Anne is tired, assists her into a carriage:

> Yes,—he had done it. She was in the carriage, and felt that he had placed her there, that his will and his hands had done it, that she owed it to his perception of her fatigue, and his resolution to give her rest. She was very much affected by the view of his disposition towards her which all these things made apparent. This little circumstance seemed the completion of all that had gone before. She understood him. He could not forgive her,—but he could not be unfeeling.
>
> (Austen, Persuasion (1997): 67)

In this passage, the voice of the narrator and the voice of Anne are unified—the narrator is telling us what Anne feels by presenting her point of view. Even more remarkable is that this point of view is one that presents Wentworth's perspective about Anne—it is a moment of insight that is central in the development of their reunion. Had this passage been presented from a purely third-person perspective the effect on the reader would have been very different: we would have been made privy to Anne's thoughts from the outside, and then we would have decided, based on the reliability of the narrative voice,

whether or not to accept the description. But here we are presented simultaneously with the narrator's perspective and with Anne's, which in this passage speak together. This dual and harmonious perspective strengthens the reader's sympathy for Anne; it also (p. 503) strengthens our belief in Wentworth's feelings for Anne. The reader is thus invited to think about Anne's thoughts from both an external and internal perspective—her views about Wentworth and the effects this has on her. We are also asked to consider Wentworth's thoughts and feelings, as well as our own assessment, as readers, of these views. This multiplicity of perspective lies at the heart of mentalization, as we seek to understand others, our views of others, their views of us, and their views of each other as well.

Ordinary life does not often present us with such opportunities for insight; via engagement with the writings of a skilled author such as Jane Austen, the reader can sharpen her mentalizing skills and develop capacities for thinking about others through a variety of perspectives, including her own. Literature thus not only requires mentalizing skills, but helps to develop and sharpen them. *Persuasion*, like many of Austen's novels, is concerned with the ways in which individuals learn to understand themselves in relation to others, and to tease out their real motivations from ones that are imposed by social convention. One of Austen's gifts is to place the reader in the centre of this journey of discovery in so subtle a way that the reader must constantly reflect on her own position within a matrix of views.

Literary *style*, then, in addition to content and narrative structure, is extremely important for engaging readers' sympathies, and formal stylistic techniques employed by a writer are able to limit or expand the ways in which readers are asked to reflect on the minds of literary characters. The more perspective-sharing we are asked to do, and the more subtly our imaginations are engaged, the greater the mentalizing capacities that are required. This is partly the case because one of the hallmarks of cognitive development is the understanding that beliefs don't merely *reflect* the world, but rather provide one of many perspectives *on* the world. When the capacity to adopt multiple perspectives is interwoven with an understanding of how perspectives interrelate and alter our understanding, both cognitive and emotional, of a character's motivations, our mentalizing capacities are strengthened. This, I am arguing, is what certain kinds of narrative styles can accomplish, especially in the hands of a gifted writer with a highly developed skill for nuance and observation.

Mentalization and Poetic Form: the Sonnet

From an examination of a narrative style, I now turn to the tightly structured form of the sonnet, and offer a close reading of one of the finest exemplars of the sonnet form— Shakespeare's Sonnet 30. Poetic form is another way in which the *formal* elements of literature enable mentalization, albeit in a way different from that of free indirect discourse. Similarities between poetry and the development of mentalization are noted by

the psychoanalyst Jeremy Holmes, but for Holmes, the focus is on psychoanalysis, not poetry. Holmes writes:

(p. 504)

> Poetry and psychotherapy are strange, yet compatible, bedfellows. Both regularly arouse suspicion and incomprehension, yet people often turn to them when in states of heightened emotion—love, elation, despair, confusion, loss and bereavement. As I have suggested, both could be seen as means with which to enhance the capacity for mentalising: the capacity to 'think about feelings' or to be 'mind-minded' … Finding the right words in the right order (Coleridge's famous definition of poetry) is a crucial skill for therapists (and their patients) as well as poets, since the appropriate image or metaphor can mirror or evoke feelings in the listener in a way that facilitates empathic attunement.

> (Holmes 2016: 43)

In exploring the connections between poetry and psychoanalytic therapy, Holmes claims that both enable 'reverie', and the 'constant interplay between feelings and language'. Both poetry and psychoanalysis enable the working-through of emotions by using language to give form to unconscious thought and feeling. Much has been said of the reparative nature of literature by Hanna Segal; of aesthetic form, she writes:

> formal modes of speech … the strictness and rigidity of the rules are all, I believe, an unconscious demonstration of the fact that order can emerge out of chaos. Without this formal harmony the depression of the audience would be aroused but not resolved. There can be no aesthetic pleasure without perfect form.

> (Segal 1952: 204)

Aesthetic form can function to order thought and emotion, and also to arouse it; for example, the limerick shares with the sonnet the feature of being highly structured. Spiller notes that the last line of the limerick combines 'conclusion with repetition' (Spiller 1992: 4) or it functions as 'the completion of the narrative, adding an extra item of information but sacrificing the echo, or reprise, of the opening line' (Spiller 1992: 5). The form of the limerick creates an expectation in the reader, who 'expects the returning rhythm of the last line to bring with it, like the fifth act of a play, a solution of previous incompleteness' (1992: 5). The brevity of the limerick, together with its use of repetition and other stylistic features, arouses in the audience amusement rather than, say, pity or fear. Another example where aesthetic form dictates a reader's (or listener's) cognitive and emotional responses is the joke: the philosopher Noël Carroll attempts to provide a theory of jokes by examining their 'underlying structural principles' (Carroll 2003: 322). He posits that the structure of the joke is strict: 'a joke is an integral unit of discourse with a marked beginning and an end' (2003: 322), and its distinguishing feature is the punch line, which 'concludes the joke with an unexpected puzzle whose solution is left to the listener to resolve' (2003: 323). Carroll draws an analogy with

Aristotle's account of tragedy: just as the formal features of a tragedy are necessary for the arousal of tragic *katharsis*, so too are the formal features of jokes in part responsible for the emotional response (laughter) to which they give rise.

The importance of form in structuring a reader's emotional and cognitive responses is evident too in the sonnet: 'compact, shapely, highly finished, and able to contain, in concentrated form, almost all that is human' (Spiller 1992: 1). The sonnet is 'probably the (p. 505) longest-lived of all poetic forms, and certainly the longest-lived of all *prescribed* forms. A prescribed form, or *closed form* … is one whose duration and shape are determined before the poet begins to write' (1992: 2, italics in original). Strictness of form primes the reader's expectations, and enables complex and often difficult emotions to be contained[8] in ways that bring insight and enable mentalization. The figurative elements of poetry—metaphor, allusion, imagery, rhyme, rhythm, and so on—contribute to the meaning of the poem via the vehicle of poetic form, which unites both form and content. The tight structure of the sonnet is able to convey the development of thought, and its brevity allows the reader to follow and keep in mind at a single reading changes in ideas and emotional affect. The sonnet form is thus in a unique position for both depicting and exemplifying the thinking process.

The sonnet form is tightly structured: in the Petrarchan sonnet, which is divided into two quatrains and a sestet, the first eight lines often act as the development of an argument, and the sestet as its conclusion. In the Elizabethan or Shakespearean sonnet, where the structure is three quatrains and a rhyming couplet, the couplet often takes on the role of presenting a 'summing up' (Spiller 1992: 4) statement. The sonnet is also well suited as a vehicle for address: 'The sonnet … always gives an impression of immediacy, as if it proceeded directly and confessionally or conversationally from the speaker' (Spiller 1992: 5) and is an ideal vehicle for introspective thought. The literary critic Helen Vendler notes that Shakespeare's sonnets are 'inward, meditative, and lyrical' (Vendler 1997: 5) rather than 'outward, expository, and narrative'. The lyric form in the sonnet 'gives us the mind alone with itself' (Vendler 1997: 19). Vendler argues that the significance of the sonnet—and of Shakespeare's sonnets in particular—does not necessarily lie in mining the *meaning* of the poem—sometimes a sonnet can be paraphrased as making such banal claims as 'I have insomnia because I am far away from you' (Vendler 1997: 13). Vendler argues rather that the 'appeal of lyric lies elsewhere than in its paraphrasable statement' (1997: 14). What, then, *does* lie at the heart of the Shakespearean sonnet? Wherein lies its value? Vendler argues that great poets use in their sonnets a variety of techniques—temporal, emotional, semantic, dramatic—to allow the development of thought within the tightly knit fourteen lines. She argues that one of the ways poetry is able to elicit emotions is by capturing the thinking process that underlies literary creation: 'Shakespeare learned to find strategies to enact feeling in form, feelings in forms, multiplying both to a superlative degree' (Vendler 1997: 17). Poetry—and the sonnet form in particular—reveals the evolution of thought in symbolic form; in poetry, thinking is 'made visible' (1997: 9). Poems may develop arguments, but unlike pure philosophical arguments, poetry appeals to emotions as well as to thoughts and ideas. Indeed, I hope to show that, in the best poems, the distinction between a thought and an

emotion is an unstable one—mentalization develops precisely when our cognitive and affective abilities work in harmony.

(p. 506) Sonnet 30

Shakespeare's famous Sonnet 30 illustrates the lyric thoughtfulness of which the sonnet form is capable.

> When to the sessions of sweet silent thought
> I summon up remembrance of things past,
> I sigh the lack of many a thing I sought,
> And with old woes new wail my dear time's waste:
> Then can I drown an eye (unused to flow)
> For precious friends hid in death's dateless night,
> And weep afresh love's long since cancelled woe,
> And moan the expense of many a vanished sight:
> Then can I grieve at grievances foregone,
> And heavily from woe to woe tell o'er
> The sad account of fore-bemoanèd moan,
> Which I new pay as if not paid before.
> But if the while I think on thee (dear friend)
> All losses are restored and sorrows end.

The poem represents within its brief length 'a multi-layered self, receding through panels of time' (Vendler 1997: 165). 'We might give such temporal panels the names "now", "recently", "before that", "yet farther back", "in the remote past"' (1997: 165). The time periods are associated with particular emotional experiences that tell of the experience of loss and mourning:

> In receding order, before the weeping 'now' (T_5, where T = Time), there was the 'recent' dry-eyed stoicism (T_4); 'before that' the frequent *be-moanèd* moan (T_3) of repeated grief; 'further back in the past,' the original loss (T_2) so often mourned; and 'in the remote past' (T_1), a time of achieved happiness, or at least neutrality, before the loss. These panels of time are laid out with respect to various lacks, grievances, and costs, as we track the emotional history of the speaker's responses to *losses* and *sorrows* (the two summarizing categories of line 14).

> (Vendler 1997: 165; italics in original)

Like all Shakespeare's sonnets, Sonnet 30 is divided into three quatrains and a rhyming couplet; each quatrain focuses on a particular idea or image, and they are then woven together to form a coherent unity. *Thought* and *thinking* form a central theme of this poem. In line 2, the present act of remembering is likened to the summoning of memories before a judge—the critic Stephen Booth notes that 'sessions' refers to 'the periodic sittings of judges, a court of law' (Booth 2000: 181). Seymour Smith 'notes that the legal metaphor "adds the notion of guilt and punishment to that of nostalgia"' (quoted in Booth: 181). The repetition of the 's' sound brings to mind the sweep of judges' (p. 507)

robes; it also serves to connect the act of memory ('summon up remembrance') with the vocal ('sighing') and the visual ('seeing')—since 'sigh' and 'sight' were often confused in Renaissance English.

It's interesting to note that the poet does not 'remember', but rather 'summon[s] up remembrances'—as though the memories are mental states distinct from the subject who remembers. This brings to mind Bion's distinction between 'thoughts' and 'thinking'—the former refers to unconscious mental states that are not subject to secondary processing and are unavailable for mentalization. It is only when we are able to *think* our thoughts that we start to develop mentalizing capacities. It is important to remember that mentalizing capacities are both inward and outward looking—the ability to mentalize refers both to our capacity to understand our own minds as well as the minds of others. When thoughts merely *occur* to the mind, the mind is experienced as passive rather than active. To reference Richard Wollheim, 'When the mind is passive, thoughts are conceived of as effecting an entry into it, from the outside' (Wollheim 1974: 35), in contrast with an active mind that regards itself as both the source and the location of thoughts.[9] As argued earlier, mentalization is a 'meta-cognitive' phenomenon, and involves thinking about thinking; a passive mind in which thoughts passively occur is unlikely to exhibit well-developed mentalizing capacities. As the narrative of the poem develops, so too does the speaker's ability to connect different aspects of his past in a way that brings forth a new understanding of his emotional life, as well as a renewed ability to mourn.

As noted in the previous section on free indirect discourse, one of the characteristics of mentalization is the ability to adopt multiple perspectives simultaneously. This ability would also include the ability to adopt hypothetical perspectives regarding one's own life. In Sonnet 30, the narrator is able to step out of the narrated time and to comment, explain, or evaluate the actions of his past self. The layered perspectives in Sonnet 30 capture well this aspect of meta-cognition. The reader of the sonnet, by being required to adopt imaginatively these various perspectives on the speaker's past selves, acquires new emotional insights in a way that real life does not often provide. A failure to engage in this imaginative work of perspective-taking will result in an aesthetic failure to appreciate the poem: the sonnet asks us to flex our mentalizing muscles via our imaginations.

Another feature of mentalization is the ability to integrate felt experiences with intellectual and emotional ones. Dimaggio et al., in a paper that examines the autobiographical memories of psychiatric patients, found that typically those patients who suffered from 'significant forms of psychopathology' (Dimaggio et al. 2012: 1) exhibited disturbances in memory: among other features, their memories lacked 'a pictorial quality which might enable a listener to imagine what happened' and memories often lack detail 'which would lend an opportunity to understand the memory as a unique experience' (2012: 2). Sonnet 30 is characterized by bringing into focus the very *situated* experiences—both emotional and somatic—of the speaker; past losses are expressed through sighing, weeping, moaning, and an eye drowning in tears. These images evoke in the reader a visceral response to the speaker's grief, which can then be processed. The thoughts about old losses bring (p. 508) forth new tears—'with old woes new wail'. The

alliteration is onomatopoeic—as is the word 'sigh' in line 3—and conjures up the physical sensations of weeping. This also emphasizes the connection between past and current expressions of grief. By engaging imaginatively with the work of mourning, the reader too embarks on an emotional journey with the speaker of the poem.

At the start of the poem, we are presented with a speaker at odds with himself and who, via the act of remembering and the activity of writing, is attempting to achieve emotional resolution. The imagery in the first two stanzas centres on the *eye* and the *mouth*, which do the work of mourning. Perhaps the visual similarity of the open mouth/eye—the 'o'—assists with the resemblance. In stanza one, the speaker does not speak but 'wail'—an infantile, preverbal expression of sadness. Stanza two is also dominated by oral and visual images: the speaker 'drown[s] an eye' for 'precious friends hid in death's dateless night'. This is a strange image of the eye drowning itself with its tears, as though enacting a kind of suicide. This suggests a strong identification by the speaker with the dead friends—since the dead friends are hidden and can no longer *be* seen, the eye ('I') will refuse *to see*. Midway through stanza two, though, the speaker 'weep[s] afresh', which suggests a move back into the present with the capacity to mourn still intact. Wailing has progressed to 'moaning', which seems a more articulate form of grieving. In stanza three, the relationship between past and present grief is again emphasized via a repetition in the language—the speaker 'grieves at grievances foregone', he moves 'from woe to woe' to 'tell o'er/The sad account of fore-bemoanèd moan', which he 'new pays' as if not 'paid before'. The repetition illustrates how emotions that were experienced in the past are being re-experienced in the present in a way that strengthens and intensifies them. It's not the case that the speaker is simply remembering past grief, but he is re-experiencing and reworking it in the present. This suggests an integration between past and present selves. Note how the emotional movement of the poem is facilitated by its structure—later quatrains are built upon the images and emotional tonality of previous ones and draw the reader forward towards the concluding couplet. Note also the change in imagery—the emphasis on bodily expression via the mouth (sighing, wailing, moaning) and the eyes (seeing, weeping) is transformed to more ego-syntonic activities—grieving and speaking ('tell o'er'). The sonnet which began with 'thoughts' ends with 'thinking'—lost emotions that were once buried and disavowed have become revived, re-experienced, and worked through so that they can be 'thought' about.

Sonnet 30 investigates the nature of grief and mourning—its vicissitudes over time, the ways in which it may transform the self and be transformed in turn. By the end of the sonnet, we note how the emotions have also transmuted into *meta*-emotions—their object is not only lost loves and deceased friends, but also the past *emotions* of grief and loss. The speaker in Sonnet 30 does not simply give expression to painful emotions, but reflects on these emotions, and in reflecting changes them. In a short fourteen lines, the sonnet expresses and re-enacts the process of mourning—from an identification with the departed loved ones in fantasy to a more mature kind of grief that has moved from an initial denial to acceptance.

Indeed, I would argue that the brevity of the poem intensifies the emotional experiences, and also requires that the reader interweaves the various time-slices of the speaker's (p. 509) multilayered self. The closed form of the sonnet asks the reader to keep in mind a range of emotional experiences and perspectives in a way that a longer poem could not do as effectively. We are induced to identify emotionally with the speaker via imaginatively identifying with his actions. We too are familiar with weeping for dead friends, and feel bereft when we relive old sorrows. The descriptions of grief in the sonnet, immersed in bodily images, have the effect of creating a visceral identification with the speaker. It is via this identification that we can engage in the working through of these emotions. This account of the reparative aspect of creativity by Hanna Segal captures this idea well:

> The reader identifies with the author through the medium of his work of art. In that way he re-experiences his own early depressive anxieties, and through identifying with the artist he experiences a successful mourning, re-establishes his own internal objects and his own internal world, and feels, therefore, re-integrated and enriched.

> (Segal 1952: 205)

Although Segal focused on the reparative aspects of creativity, I think her general insights also help explain how literary and creative artworks facilitate mentalization. Sonnet 30 both describes and enacts an experience of mentalization, in part via encouraging the emotional identification with the dramatic speaker to which Segal refers. This ability to see from the perspective of another forms a central aspect of mentalizing capacities; the ability to then reflect further upon these various perspectives and subject them to re-examination contributes to the development of higher-order cognitive and emotional capacities.

Some Final Comments

The examples discussed in this chapter of the ways in which formal features of literature, such as narrative style and sonnet form, may facilitate mentalizing capacities are clearly not exhaustive. This chapter takes a small but hopefully important step in making the case for the importance of *stylistic* features of artworks, as opposed to focusing only on their narrative content. An implication of this claim is that different genres will facilitate different kinds of emotional and cognitive responses and meta-responses. I have focused to a large extent on the ways in which formal literary features that encourage *perspective taking* facilitate mentalization. Although I think that the ability to adopt multiple perspectives is an important feature of mentalization, I do not want to suggest that it is the only one. For example, other stylistic features of poetry, such as rhyme, rhythm, and imagery, may evoke emotional responses in particular ways: these will then contribute to the overall experience of the literary work, sometimes by working harmoniously together, sometimes by creating tensions that require resolution. The ways in which stylistic, formal, and narrative elements work together or pull apart forms part of the complexity

of literary works and our engagement with them. In 'Creative Writers and Day-Dreaming', Freud drew parallels between the work of the creative writer and (p. 510) the child at play —both create 'a world of phantasy' which they invest with 'large amounts of emotion— while separating it sharply from reality' (Freud 1908: 144). On this view, the purpose of creative writing is the expression of desire—often unfulfilled—via the imagination. There is no doubt much truth to this comparison; but where the creative writer of good literary fiction differs from the mere teller of tales is by tying this expression of fantasy to the deeper psychological work of attaining insight. The literary writer crafts the expression of unconscious fantasy in a way that transforms it so that it can be understood. In the abstract to this chapter I wrote that book reading shares some of the same features as mind reading: reading (good) literature enables us to develop—if partially—the same kinds of mentalizing capacities as we acquire in therapy or analysis. When we read, we engage in a kind of conversation, which shares some features of the 'peculiar conversation'[10] that is psychoanalysis. This view of both literature and psychoanalysis places them at the heart of a humanistic endeavour—reading the great writers makes us more human because we share in *their* insightful humanity. Reading—like analysis—is at heart relational; our minds develop in consort with others.

References

Austen, J. (1997). *Persuasion*. New York: Dover Publications (Original work published 1817).

Booth, W. (1983). *The Rhetoric of Fiction*, 2nd edn. Chicago, IL: University of Chicago Press.

Booth, S. (2000). *Shakespeare's Sonnets*. New Haven, CT/London: Yale University Press.

Bray, J. (2007). 'The "dual voice" of free indirect discourse: a reading experiment'. *Language and Literature 16*(1): 37–52.

Carroll, N. (2003). 'On jokes'. In *Beyond Aesthetics* (pp. 317–335). Cambridge: Cambridge University Press.

Chen, S. (2014). 'Austen's narrative perspective and the problem of interpretation in *Emma* and *Persuasion*'. *Tulane Undergraduate Research Journal 1*: 32–39.

Cohn, D. (1966). 'Narrated monologue: definition of a fictional style'. *Comparative Literature 18*(2): 97–112.

Currie, G. and Ravenscroft, I. (2002). *Recreative Minds*. Oxford: Oxford University Press.

De Maupassant, G. (1956). 'The jewels'. In *88 Short Stories* (pp. 370–377). London: Cassell & Co Ltd.

Diamond, D. and Kernberg, O. (2008). 'Discussion'. In F. N. Busch (ed.), *Mentalization: Theoretical Considerations, Research Findings, and Clinical Implications* (pp. 235–260). New York/Sussex: Analytic Press.

Dimaggio, G., Salvatore, G., Popolo, R., and Lysaker, P. H. (2012). 'Autobiographical memory and mentalizing impairment in personality disorders and schizophrenia: clinical and research implications'. *Frontiers in Psychology 3*: 1–4.

Feagin, S. (1983). 'The pleasures of tragedy'. *American Philosophical Quarterly 20*(1): 95–104.

Fonagy, P. (2008). 'The mentalization-focused approach to social development'. In F. N. Busch (ed.), *Mentalization: Theoretical Considerations, Research Findings, and Clinical Implications* (pp. 3–56). New York/Sussex: Analytic Press. (p. 511)

Fonagy, P., Gergely, G., Jurist, E., and Target, M. (2002). *Affect Regulation, Mentalization, and the Development of the Self*. New York: Other Press.

Freud, S. (1908). 'Creative writers and day-dreaming'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 9 (pp. 143–153). London: Hogarth Press.

Hjort, M. and Laver, S. (1997). *Emotion and the Arts*. New York: Oxford University Press.

Holmes, J. (2005). 'Notes on mentalizing—old hat, or new wine?' *British Journal of Psychotherapy 22*: 179–197.

Holmes, J. (2016). *The Therapeutic Imagination: Using Literature to Deepen Psychodynamic Understanding and Enhance Empathy*. Oxford: Routledge.

Jurist, E. L. (2005). 'Mentalized affectivity'. *Psychoanalytic Psychology 22*(3): 426–444.

Kidd, D. C. and Castano, E. (2013). 'Reading literary fiction improves theory of mind'. *Science 342*: 377–380.

Lear, J. (2005). *Freud*. Oxford/New York: Routledge.

Mar, R., Oatley, K., and Peterson, J. (2009). 'Exploring the link between reading fiction and empathy: ruling out individual differences and examining outcomes'. *Communications 34*(4): 407–428.

Nichols, S. (2004). 'Imagining and believing: the promise of a single code'. *Journal of Aesthetics and Art Criticism 62*(2): 129–139.

Robinson, J. (2005). *Deeper than Reason: Emotion and its Role in Literature, Music, and Art*. Oxford: Clarendon Press.

Segal, H. (1952). 'A psycho-analytical approach to aesthetics'. *The International Journal of Psychoanalysis 33*: 196–207.

Segal, H. (2000). *Dream, Phantasy and Art*. Hove/New York: Brunner-Routledge.

Spiller, M. R. G. (1992). *The Development of the Sonnet: An Introduction*. London: Routledge.

Target, M. (2008). 'Commentary'. In F. N. Busch (ed.), *Mentalization: Theoretical Considerations, Research Findings, and Clinical Implications* (pp. 261–279). New York/Sussex: Analytic Press.

Vendler, H. (1997). *The Art of Shakespeare's Sonnets*. Cambridge, MA: Harvard University Press.

Wollheim, R. (1974). *On Art and the Mind*. Cambridge, MA: Harvard University Press.

Wood, J. (2008). *How Fiction Works*. London: Jonathan Cape.

Woolf, V. (1996). *To the Lighthouse*. London: Penguin Books (Original work published 1927). (p. 512)

Notes:

(¹) A great deal of contemporary literary criticism that explores the themes of novels, plays, and poetry via the insights provided by psychoanalysis owes much to the French psychoanalytic school, which is not my focus.

(²) For a fuller discussion on film and dreams, see Cox and Levine (this volume).

(³) See, for example, Kidd and Castano (2013), and Mar, Oatley, and Peterson (2009).

(⁴) See, for example, Currie and Ravenscroft (2002); Nichols (2004); and Hjort and Laver (1997).

(⁵) I am grateful to Michael Lacewing for this observation.

(⁶) I would like to thank Richard Gipps for this observation.

(⁷) This seems borne out by empirical research: in an experiment designed to test whether free indirect discourse requires that readers shift their points of view while reading, the researchers conclude that in some uses of free indirect discourse, 'the ambiguity may never be resolved and both points of view may remain in play' (Bray 2007: 48).

(⁸) Hanna Segal speaks of the importance of form in the containment of feeling with regard to tragedy. Tragic form 'contains feelings which otherwise might be uncontainable' (Segal 2000: 90).

(⁹) See Freeman (this volume) for a fuller discussion on Wollheim.

(^{10}) I borrow this phrase from Jonathan Lear, which he uses as a title to the Introduction to his book *Freud* (2005).

Elisa Galgut

Department of Philosophy, University of Cape Town, South Africa

Oxford Handbooks Online

Abstract and Keywords

Those who believe that the psychoanalytic understanding of human nature is broadly correct will also likely believe that there are essential aspects of film that cannot be adequately understood without it. Among these are film's power; the nature of film spectatorship; and the characteristics of specific films and genres. Why are we attracted to certain kinds of films—horror films and those depicting violence we abhor? The most basic claim underlying psychoanalytic approaches to film is that the creation and experience of film is driven by desire and wish-fulfilment and functions to satisfy certain psychological, protective, expressive needs of artists and audiences. Psychoanalytic explorations of film tend to draw together aspects of artistic creation and spectatorship, as well as accounts of film's power to move audiences and the nature of film spectatorship in general—the affective and cognitive significance of the nature of film experience itself.

Keywords: spectatorship, power, horror, wish-fulfilment, desire, violence, affective and cognitive significance, protective and expressive needs, human nature, experience

Damian Cox and Michael Levine

(p. 513) Introduction

PSYCHOANALYSIS is both a set of theories of the structure and workings of the mind and the psychotherapeutic method based on Freud's theory of the mind. Those who believe that the psychoanalytic understanding of human nature and the mind are broadly correct will also likely believe that there are essential aspects of film that cannot be adequately understood without it.[1] Prominent among these are (i) film's power; (ii) the nature of film spectatorship; and (iii) the characteristics of specific films and genres. Why are we attracted to certain kinds of films, for example horror films and those depicting violence

we abhor?[2] Psychoanalytic approaches to film tend to be more unified than piecemeal approaches that see issues of spectatorship and narrative construction and genre as discrete questions. However, the fact that psychoanalysis approaches film with a unified theory of mind, though there is considerable dispute as to its details, should not be confused with the claim that psychoanalytic approaches are exhaustive and complete.

(p. 514) Richard Allen (2009: 446) says:

> Orthodox psychoanalytic theories of art have focused on [1] the relationship between the creation of art, sexuality, and unconscious mental life, and a great deal of both film and literary criticism has used psychoanalytic theory to interpret texts or genres of texts, but the distinctive contribution of film theory to psychoanalytic theories of art lies in [2] its focus upon the nature and character of film spectatorship.

From a psychoanalytic perspective [1] and [2]are closely related. Freud sought to explain not just the meaning of texts and their relation to the artist—what the artist is doing and how art functions both for artists and audiences. He also sought to explain why audiences took an interest in such works, and how and why art is able to convey phantasies, wishes, and desires in ways palatable to audiences.[3] The most basic claim underlying psychoanalytic approaches to film is that the creation and experience of film is driven by desire and wish-fulfilment and functions so as to satisfy certain psychological, protective, expressive needs of both artists and audiences.

Psychoanalytic explorations of film tend to draw together aspects of artistic creation and spectatorship. The conditions of a work's creation and the ways in which it is viewed are connected, though not identical. Psychoanalytic explorations of film also tend to draw together accounts of film's power to move audiences and the nature of film spectatorship in general (the affective and cognitive significance of the nature of film experience itself). They also tend to draw together accounts of film's power to move audiences and the characteristics of particular genres of films—such as horror films or revenge films.

When it comes to film theory, psychoanalytic influence has been profound. Christian Metz's Lacanian account of film spectatorship and Laura Mulvey's account of the structure of cinema as grounded in male voyeurism are two very prominent and influential psychoanalytic interventions in film theory, but the interactions between the two fields has been long-standing, extensive, varied, and deep.[4] In the late 1980s, however, a group of film theorists overtly hostile to psychoanalytic ways of approaching film came to prominence. Film studies undertook a cognitivist turn, and one of the foundational cruxes of cognitivism was a rejection of the psychoanalytic modes of explanation.[5] Cognitivists in film theory attempt to explain basic features of film experience: audience comprehension, emotional elicitation, character identification, and aesthetic preference.[6] They do so using the resources of contemporary cognitive science and analytic philosophy, (p. 515) self-consciously eschewing the work of psychoanalytic theorists. They represent the most trenchant critics of psychoanalytic approaches to the study of film. In our view, the rejection was premature and exaggerated. We discuss

cognitivism and its rejection of psychoanalytic approaches to film later in the chapter, but first we offer a clarification of applications of psychoanalysis. We then explore the positive side of the story. What relevance might psychoanalysis have to the creation and experience of film art?

Applying Psychoanalysis to Film

Psychoanalytic aesthetics has drawn from and contributed to psychoanalytic theory. Freud and psychoanalytic theorists generally have seen art as supporting and enlarging such theory. For those who accept the basic tenets of psychoanalysis, its engagement with aesthetics has advanced theory by enhancing the psychoanalytic understandings of mind and human nature.[7] Basic tenets of psychoanalysis that are especially relevant in this context include such things as the reality of dynamic unconsciousness, the reality of 'primitive' (primary process) mental function focused on immediate gratification of instincts and drives, the centrality of defence mechanisms for managing anxiety, the emotional meaningfulness of dreams and phantasy. Some features of film spectatorship are best explained by our tendency to repress and disavow uncomfortable or perverse desires, and to gain a special kind of gratification from the vicarious satisfaction of these desires in dreams and phantasies. In dreams and phantasies, representations beset us and our ordinary modes of self-critical judgement are suspended. Much the same thing seems to happen, at least sometimes, in film spectatorship.

Much as Freud used what he termed 'the psychopathology of everyday life', film and other arts are used to support the validity of psychoanalysis and theoretically to enhance it.[8] An understanding of art on the one hand and psychoanalysis on the other is mutually supportive. An understanding of both art and psychoanalysis is furthered by explaining fundamental features of aesthetics: its significance in virtue of the roles that it plays in our lives; the relation between artists and their work, as well as between artists and audiences; and even why there is any art at all.

It is a common misconception of psychoanalytic approaches to film and art generally to think that if there are exceptions to Freud's claim that art is driven by desire and wish-fulfilment, and functions to satisfy certain psychological, protective, expressive (or whatever) needs of artists and audiences, then one will have shown Freud's views (p. 516) about art to be mistaken.[9] Clearly, exceptions undermine claims of universal and necessary truth, but psychoanalytic approaches to film need not make such claims. Wood (2004: xv) expresses the point this way.

> Part of the problem lies in that distressingly common tendency either to totally accept or totally reject, as opposed to the principle of *examining critically*. Few today appear to read Freud or Marx with a view to sorting out what is still valid, what can be cast off, and what needs to be rethought. Freudian theory is vulnerable to attack on many points, but not, in my opinion, on the one that formed *The American Nightmare*'s psychoanalytic basis: the theory of repression and the 'return of the repressed.' We can all trace the workings of this, surely, in

our own personal histories and daily lives; it continues to have great resonance in relation to the horror film . . .

Wood may overemphasize the extent to which the mere unburdening of repression and the 'return of the repressed' is at work in film spectatorship. Repression is just one form of defence explored in psychoanalytic theory. Nonetheless, Wood makes clear the idea that psychoanalytic insights into film are not an all-or-nothing, take-it-or-leave-it phenomenon. Getting clear about the way psychoanalysis may, on the one hand, share the stage with other interpretative methods and tools and, on the other, characterize some fairly general aspects of the spectatorship of genres (such as horror films) will be a task for our discussion of psychoanalysis and genre later. For now, it is worth pointing out the basic fallacy of rejecting any psychoanalytic theorizing about film or interpreting films and film genres by problematizing its blanket, universal application. For example, certain films—say a documentary on an economic crisis such as Craig Ferguson's *Inside Job* (2010)—seem neither to invite nor reward psychoanalytic treatment (as opposed to the people depicted in the film, who surely do invite such treatment).

Prominent uses of psychoanalysis in film theory throughout the 1960s and 1970s tended to offer universal and fundamental analyses of the experience of film. Christian Metz (1982), for example, emphasized the 'gaze' and its psychoanalytic significance. The idea is that the appeal of watching meaningful, moving images on a screen is based on an identification with the camera and corresponding fetishistic scopophilia. This is an essentializing move: a way of explaining the key features of all cinematic experience within a single theoretical perspective. This essentializing approach reached its apotheosis with 'apparatus theory' in the 1970s. Apparatus theory combined psychoanalytic, semiotic, and Marxist theoretical perspectives in one theory—one that made the psychoanalytic and ideological character of film spectatorship an inevitable product of the physical and institutional apparatus of film production (see Baudry 1986).

The attempt to construct an essentialist theory of cinematic spectatorship, one that characterizes spectatorship of the moving image as such within one overarching theory, has largely been abandoned. This leaves open the possibility that psychoanalytic insights are an important and ineliminable framework to explain particular aspects of film (p. 517) spectatorship. Some writers, however, remain unconvinced of even this.[10] For example, Richard Allen is largely critical of psychoanalytic approaches to film. He describes '[p]sychoanalytic theory [as] a theory of the relationship between sexuality and unconscious mental states … ' (2009: 446). Part of the reason Allen is critical of psychoanalytic approaches to film is that he mischaracterizes and oversimplifies psychoanalytic theory. It is true that psychoanalysis is concerned with sexuality and unconscious mental states, but it is also misleading as a characterization generally, and especially as one that can adequately ground the relevance of psychoanalytic theory to film. It mirrors the untutored idea, along with objections based on that idea, that psychoanalysis is *all about sex*. Not only does sexuality have to be interpreted broadly with reference to bodily and mental pleasures (not sex as ordinarily understood); but connections between sexuality and unconscious mental states have to be understood in

relation to the orectic (i.e. to desire and appetite). If this is not done, then it is difficult to envision psychoanalytic film theory as even getting started.

Gabbard (2001: 5–12), the first film review editor for the *International Journal of Psychoanalysis*, gives an exhaustive overview of psychoanalytic approaches to film, including the explication of underlying cultural mythology, the film as reflective of the film-maker's subjectivity, the film as reflective of a universal developmental moment or crisis, the application of Freud's dreamwork to film, the analysis of spectatorship, the appropriation of psychoanalytic constructs by the film-maker, and the analysis of character in the narrative. These overlap and are by no means mutually exclusive. Analysis of some films may require several approaches, and it should go without saying that psychoanalytic interpreters of film do not see themselves, and should not be seen, as employing the only useful approach to understanding character, narrative, or spectatorship—not even when such an approach is thought of as necessary. As Gabbard (2001: 12–14) says:

> Some psychoanalytic film critics deliberately mix methodologies for a more comprehensive reading of a particular text … [When] A multiplicity of theoretical perspectives is brought to bear … the result is a psychoanalytic film criticism that understands a particular movie as having multi-layered meanings that are not immediately apparent to the average viewer.

Nor is it plausible to suppose that the perspectives are all psychoanalytic.

Much of the controversy in applied analysis has revolved around whether the appropriate subject for analysis is the art object itself or, rather, the biographical features of the artist that may contribute to our understanding of the forces shaping the artistic creation. Both may be fruitful subjects for exploration, and psychoanalytic film scholars have (p. 518) made productive use of both approaches. Obviously, when one applies a psychoanalytic lens to the text of a film, one cannot hope for a definitive reading. A more modest goal is to emphasize how psychoanalytic theory can often illuminate what appears to be happening on the screen and the manner in which the audience experiences it.

To sum up this discussion: psychoanalytic accounts of the nature and function of art and film should not be taken, as they often are by their critics, as reductive. Psychoanalytic accounts do not rule out other purposes, functions, and explanations of art. These include the rational exploration of ideas and feelings as entertaining, as providing cognitive insights of various kinds, as morally examining the social, political, and personal status quo.

The Orectic Self: Psychoanalysis, Dreams, Film

Psychoanalysis claims that psychic life is dominated by our orectic nature—that is, driven by desire and wish-fulfilment. We are not the rational, transparent, self-knowing creatures we generally take ourselves to be. Art too is orectic. It functions, though not

exclusively, to satisfy psychological needs of artists and audiences, and relates artists to the audiences in virtue of common desires and satisfying phantasies. Gabbard (2001: 5–6) for example says:

> Just as dreams function as wish-fulfilments (at least in many cases), so do films provide wish-fulfilling solutions to human dilemmas. In the 1946 classic, Frank Capra's *It's a Wonderful Life*, for example, the post-World War II audience vicariously experienced the magical resolution of three pervasive internal conflicts: adventure/domesticity, individual/community and worldly success/ ordinary life … the film does not so much resolve these anxieties as push them to one side … Audiences loved Capra's film, however, because its ending so completely disposed of what had briefly returned from the repressed.

Psychoanalysis sees these needs, desires, and phantasies, along with the artist's often unconscious intentions, as necessary for interpreting film and for understanding spectatorship (e.g. spectator satisfaction). Such a view also provides grounds for psychoanalysis to see art as integral and necessary to living well.

Films function psychoanalytically in various other and related ways as well. Gabbard (2001: 7) describes some films as 'reflective of the film-maker's subjectivity … [as] a canvas in which the director attempts to work through and repair problematic childhood experience and conflicts'. And again drawing on the likeness between dreams and film, he says (2001: 8) that 'Certain films defy conventional analysis and understanding unless they are viewed as dreams subject to condensation, displacement and other elements of Freud's dreamwork'. Many of the films of Alfred Hitchcock are best understood in this way. For example, *Vertigo* (1958) casts its protagonist into a dream-like following of a phantasy (p. 519) object: its loss, its replacement, and final loss—at which exact point the protagonist's vertigo is cured. It is difficult to see how the film can be understood at all without deploying psychoanalytic categories of explanation.

The alleged connections and likenesses between dreams and film is often emphasized by psychoanalytic approaches and many are summed up in Allen's (2009: 448–9) account of McGinn (2005).

> McGinn has explored the film–dream analogy … plausibly noting the way that our familiarity with dreams tutors our experience of film … Films, like dreams, are characterized by sensory/ affective fusion … like dreams, they are characterized by spatial discontinuity and by temporal fixation …. are attention dependent … are often characterized by a heightened sensation of movement that is linked to the solicitation of strong emotion … are characterized by the 'salience' of every element, at once compressing information and amplifying emotional impact … [M]ost contentiously … McGinn claims that in films, like dreams, the minds of others seem peculiarly transparent to the spectator. The body or face in a dream is designed to express a given mind, in this sense it is a transparent portal to the mind in a way that the face of the other usually fails to be. Likewise in films, the

human mind is not merely inferred from bodily criteria, but appears something we have transparent access to.

Allen (2009: 450) also cites ' … many disanalogies between dreams and watching a movie … the most important of which is that the spectator is wide awake and the images are real (Metz 1982). The analogy … at best provides a partial understanding of the film experience and an understanding that is probably not best cashed out in psychoanalytic terms'.

The analogies between dreams and film spectatorship are nonetheless salient to understanding film spectatorship. Moreover, psychoanalytic comparison between dreams and film spectatorship is not pointing to a mere analogy between them, but to a common explanation of them in terms of psychic functioning. Most obvious here is the psychic defence against anxiety and corresponding wish-fulfilment of both dreams and film spectatorship.

A Case Study: Science Fiction Disaster Films

Let us consider an example of this: science fiction disaster films. In her seminal essay 'The Imagination of Disaster', Susan Sontag (1965) argues that science fiction films from the 1950s and 1960s (significantly, during the height of the Cold War), and viewers' attraction to them, should be seen as reflecting a widely held fear of catastrophe as well as ambivalence towards science, scientists, technology, and politicians. The films display anxiety and worry about disaster, nuclear annihilation, and the like, while at (p. 520) the same time providing easy wish-fulfilling answers to such anxieties.[11] A basic claim of psychoanalytic theory is that dreams function in part to protect sleep. If this is right, there is a further analogy to be drawn between dreaming and film. The kinds of science fiction films Sontag discusses function so as to enable people to avoid confronting and consciously examining their anxiety, as well as consciously considering strategies for dealing with real problems.[12] They function, as it were, to keep us asleep.

It is no surprise that movies (some movies) have long been seen as forms of escapism— including ones that do not, on the surface, appear as escapist. The films Sontag discusses, on her own account, really do aid and abet moral and existential escapism. Lest this sounds overly negative, it should be remembered that psychoanalytically speaking escapism of this kind is *at times* as necessary as it is desirable. Such escapism is generally desirable only temporarily as the real world will not indefinitely be put off. However, it is a mistake to think that one should or could permanently escape escapism, just as it is a mistake to think that, psychoanalytically speaking, complete self-transparency, or a life completely free of neuroses, is possible.

A good deal of what Sontag discusses has to do with catastrophe on a global scale; where what is being considered is the obliteration of earth or life as we know it on earth. One thing that Sontag's essay overlooks—perhaps because the films themselves do not deal with it, and given the way they function, could not deal with it—is the aftermath of

catastrophe. That is, the rebuilding following extraterrestrial invasion and devastation. What happens when the monsters are gone? The reason for such oversight is clear. Such films function, after all, to relieve rather than enhance anxiety and questions about what is to be done in the aftermath are surely anxiety producing.

Sontag says (1965: 224) 'From a psychological point of view, the imagination of disaster does not greatly differ from one period in history to another. But from a political and moral point of view, it does.' On this view, the imagination of disaster does not change because psychologically speaking, neither the precipitating concerns and fears (death, loss of love, meaninglessness), nor the ways in which people's minds endeavour to assuage them, substantively differ from disaster to disaster.[13]

(p. 521) Although science fiction films are 'strongly moralistic' (1965: 216), Sontag notes that they contain 'absolutely no social criticism of even the most implicit kind … No criticism … of the conditions … which create the impersonality and dehumanization which science fiction fantasies displace onto the influence of an alien It' (1965: 223). Why the lack of critical concern? Sontag's account of the function of these films carries an implied explanation. Serious social criticism is not merely beside the point, but would also prevent the films from providing the satisfactions audiences seek. Though the satisfactions are largely psychological, they may serve to emotionally ground specific beliefs and, more broadly, one's cognitive outlook—one's understanding of reality—on the social, political, and personal status quo. By their very nature, the science fiction films Sontag discusses cannot concern themselves with serious social or political criticism, even though they may express it. Any serious questioning of the moral and political status quo—conditions that are responsible for the disasters befalling people—would hamper the operation of phantasy and its production of temporarily satisfying 'solutions' to whatever catastrophe is being depicted.[14] Of course, it is possible that science fiction films offer serious social and philosophical reflection. Andrei Tarkovsky's *Stalker* (1979) is a philosophically rich post-disaster film. It follows three men—a scientist, a writer, and their guide, the stalker of the title—as they plot their way through a forbidden zone, the site of a mysterious disaster. However, films such as this are not typical of the science fiction genre and do not generally capture the attention of science fiction fans. Tarkovsky's film is first and foremost an art film, not a science fiction film.

Cognitivist Rejection of Psychoanalytic Film Theory

As noted in the introduction, cognitivists attempt to explain basic features of film experience using the resources of contemporary cognitive science and analytic philosophy, self-consciously eschewing the work of psychoanalytic theorists. Without recourse to psychoanalytic concepts and ideas, they attempt to explain such things as the identification audiences succeed in making with film protagonists; how audiences are

able to follow the plot of films, given the gaps in narrative and shifting perspectives characteristic of most films; and whether the moving image is an illusion or not.

The cognitivist film theorist Noël Carroll offers one of the most sustained and trenchant criticisms of psychoanalytic approaches to film (1988, 1990, 1996, 2004b). His championing of a cognitivist turn in film studies has been so successful that such approaches have come to be seen as incompatible with psychoanalytic approaches. **(p. 522)** However, this is only the view from the cognitivist side. The key questions of cognitivist film theory differ from, but do not replace, the key questions of psychoanalytic approaches to film. In this section, we set out what is at issue in the debate between cognitivist film theorists and psychoanalytically informed film theorists.

There are two key psychoanalytic claims that cognitivists by and large take issue with. First, things such as narrative curiosity are not primarily, let alone exclusively, capable of explaining the pleasures of spectatorship. Second, these pleasures require psychoanalytic interpretation. *If* the psychoanalytic account of the mind and psychic life is broadly correct, then its interpretation of the pleasures of cinematic spectatorship is broadly correct too. On a psychoanalytic account it is, for example, no more possible to explain prejudices as cognitive mistakes (e.g. 'people of colour are lazy; Jews are lascivious'), rather than in terms of ego defence, than it is to explain the attraction to horror film in cognitive terms while ignoring our orectic natures.

Cognitivist criticism of psychoanalytic approaches to film is based partly on scepticism about the credentials of psychoanalysis as a theory of mind and partly on a mistaken essentialism. We discuss this scepticism later, but first let us consider the essentialist presupposition of the cognitivists. We mentioned earlier certain essentializing forms psychoanalytic film theory has taken in the past. We claimed that this tendency to produce an essentializing theory of the cinema reached its most extreme form in 'apparatus theory', the theory prominently advanced by Jean-Louis Baudry and based on the claim that the apparatus of cinema—the means of the production of cinematic experience—is essentially a psychoanalytically explicable process. Interpreted as a totalizing account of the possibility of cinematic experience, apparatus theory suffers from theoretical overreach. However, it seems plausible to say that the cognitivists' mistake is to over-react to this and substitute an overreach of their own. Psychoanalytical insights into cinematic experience are not limited to grand theories such as apparatus theory; and if psychoanalytic accounts of the orectic nature of our psychic lives, in the cinema and out of it, have any merit, then we should expect highly context-laden and partial explanations of cinematic experience to emerge from it. Cinematic experience, its pleasures, and its dreads, will necessarily depend upon the peculiarities of the subject and the subject's psychological history and character. One-size-fits-all theories of cinematic experience run counter to the fundamental claims of psychoanalytic theory.

However, this does not rule out the possibility of quite general claims being made about genre and the nature of the filmic experiences elicited by a genre. Genres are characterized by a commonality of experience as well as structure and narrative.

Typically, we go to a musical film (whether *La La Land* (2016) or *Meet Me in St Louis* (1944)) for a certain kind of experience: musical pleasure, escapist entertainment, the aesthetic joy of watching people who can dance, dance. The goal, therefore, of psychoanalytic explorations of genre is to understand the typical experiences of typical audience members in typical films of the genre. There is a lot of room in this for exceptions of all kinds; and it is compatible with a realization that there may be not one kind of typical audience member or not one kind of typical filmic experience of genre. Proffered psychoanalytic explanations are not essentializing about genre, but are fairly general in their explanatory ambitions nonetheless.

(p. 523) For several reasons, the best genre for contrasting psychoanalytic and cognitivist approaches to cinema is the horror genre. On the one hand, it presents an apparent paradox—the so-called paradox of horror—which is fertile ground for psychoanalytic insight. (The paradox of horror is simply the question of why horror films generate audience pleasure from negative emotions of fear and disgust.) The horror genre appears highly suited to psychoanalytical explanation as, for a psychoanalytically orientated film theorist, the horror genre presents a striking insight into the nature of cinematic phantasy. On the other hand, the genre has been the site of robust cognitivist attempts to do away with psychoanalytic explanation and replace it with narrative-based explanations. (Very roughly: audiences enjoy watching horror films because they want to know what happens when a monster is let loose. Why? Because they find the idea of it interesting; their curiosity has been piqued.) If the cognitivist rejection of psychoanalytic approaches to film is to gain real purchase, it will be in virtue of the fact that the cognitivists can show psychoanalytic explanations of the experience of horror films to be either wrong-headed or entirely superfluous.[15]

Cognitivist and Psychoanalytic Accounts of Horror

Noël Carroll is the chief architect of this cognitivist take on horror film experience.[16] Carroll's account of the (ir)relevance of psychoanalysis to horror films contrasts very nicely with Schneider's (2000) account. Carroll (2004b: 263) says:

> Schneider points out that in Freud's [1919] characterization of the uncanny, which he finds useful for modeling horror, Freud indicates that not only repressed wishes, but also [the reconfirmation of] surmounted beliefs, can function to trigger the sense of uncanniness ([Schneider] 2000: 172). These surmounted beliefs include things like infantile beliefs in the omnipotence of the will and the belief that the dead can return to life.

The 'return of the repressed' temporarily satisfies certain wishes and desires and, according to Schneider, partly explains our attraction to horror.[17]

Psychoanalysis and Film

Schneider also discusses an aspect of the uncanny in relation to horror—a favourite of Alfred Hitchcock's—known as 'doubling'.[18] Perhaps the most striking example of doubling in Hitchcock's oeuvre occurs in *Shadow of a Doubt* (1943). In that film, Hitchcock tells

(p. 524) the story of two Charlies: a killer, played by Joseph Cotton, and his favourite niece, Charlotte (Charlie), played by Teresa Wright. The two play off each other, with parallel and opposing narrative paths. The effect lends a sense of the uncanny to a story that would otherwise be fairly routine.

Carroll concedes the fact that psychoanalysis might be interpretatively relevant because it influences the way in which artists and audiences frame their self-understanding. He calls this a hermeneutical defence of psychoanalytic interpretation.[19] However, this hermeneutical defence does not go very far. He writes (2004b: 259; cf. Carroll 1990):

> This defense of the relevance of psychoanalysis to the horror film, however, does not entail that psychoanalysis is relevant to the interpretation of *all* horror films. For, pervasive though psychoanalytic thought may be, it is not the case that in every horror film one will find evidence of psychoanalytic concepts, scenarios, and/or imagery.

And again Carroll (2004b: 260–1) says 'I agree that *many* horror films deserve a psychoanalytic interpretation (at least in part), [but] they [psychoanalytic critics] believe that *all* horror films should be interpreted psychoanalytically'. But there is no reason to insist that psychoanalysis is relevant to *all* horror films. As we have emphasized previously, it is enough if psychoanalysis is relevant to typical films of the genre and typical genre-audiences.

The reason those in favour of a psychoanalytic approach would reject Carroll's 'hermeneutical defence' is not because they think all horror films require psychoanalytic interpretation.[20] They reject it because as characterized by Carroll it is beside the point for the kind of psychological explanations of spectatorship and interpretation of films that psychoanalytically informed theorists and interpreters seek. Perhaps Carroll would agree. He says (2004b: 260–1):

> One reason why many psychoanalytic critics are apt to reject what I've called the hermeneutical defense of their practice is that they believe that psychoanalysis is true, (p. 525) whereas the preceding hermeneutical defense does not require that psychoanalysis be any less wacky than scientology to be apposite in a given case.[21]

This gets closer to why psychoanalytically orientated critics reject Carroll's hermeneutic defence. Carroll's disdain is palpable and it is clear that the rejection of psychoanalytic approaches to film frequently rests on a wholesale rejection of psychoanalysis itself.[22]

Carroll (2004b: 261) says:

. . . the case for explicating all horror films psychoanalytically will depend on showing that there is something in the very nature of the horror film that is peculiarly suited for psychoanalysis. That is, the psychoanalyst needs to establish that there is something in the essence of horror—something without which a film would not be a horror film—that is only explicable or that is best explained psychoanalytically.

Yet one can easily drop the claim that horror film 'is only [and exclusively] explicable' in psychoanalytic terms, and still claim that horror films are largely 'best explained psychoanalytically'.

Tudor (1997: 449) challenges just this claim. The psychoanalytic explanation of the pleasures associated with horror in terms of the return of the repressed, e.g. as offered by Schneider, is inadequate. He says that 'it is necessary to pose supplementary mechanisms to bridge the gap between a general account of repression and the specific explanation of pleasure, and these supplementary mechanisms lead away from the pure form of the repression model'. But such supplements will be part of any more complete explanation of the attraction of horror in terms of the return of the repressed. Unsurprisingly, there are various and often conflicting accounts of what additional elements are needed to explain the attraction of horror. Thus, Kristeva's notion of 'abjection', taken up by Barbara Creed in her (1993) theorization of the 'monstrous-feminine' in terms of a Lacanian account of the 'Real' and the construction of the feminine as 'Other', seeks to give a more complete explanation of the attraction of horror.[23] The extent to which these notions, and Lacanian psychoanalysis generally, are compatible with what Freud says about horror and Freudian theory in general, is controversial. But psychoanalytically informed (p. 526) elaborations on the return of the repressed, if evidentially defensible, support rather than subvert this model.

Although Tudor regards recourse to additional explanation in terms of psychoanalytic theory as ad hoc, it need not be. In psychoanalysis as elsewhere, more complete explanations require additional detail. The return of the repressed may be pleasurable for a variety of reasons depending on the repressed element and depending also on the particular spectator. The pleasures of horror, dependent as they are upon the effects of the repressed, may involve substitutive satisfactions—much like neurotic activity which provides replacement satisfaction for something that did not occur.[24]

Carroll (2004b: 264) appears to back away from his claim that psychoanalysis is not important to interpreting some films and aspects of spectatorship beyond the false friend of his hermeneutical defence. He says:

> that psychoanalysis might not be applicable to all horror films does not entail that it cannot clarify some. Psychoanalysis is a vast and complex body of ideas, including not only meta-psychological theories but also observations of hithertofore scarcely noticed patterns of human behavior. Some of these ideas, if

they are well founded and if they track the phenomena on the screen, may illuminate otherwise perplexing aspects of particular horror films.

Carroll's rejection of psychoanalytic approaches dwindles here to the question of just how illuminating, relevant, and sometimes necessary psychoanalytic approaches are—rather than whether they are relevant at all. He appears willing to accept some psychoanalytic observations and insights, but not the theory that explains or supports them. In any case, the concession is short-lived. He continues (Carroll 2004b: 266):

> My own suspicion … is that where most psychoanalytic interpretations of horror films succeed, it is on historicist grounds, but that may only reflect my skepticism about how many well-founded, uniquely psychoanalytic observations there are to be had about recurring patterns of human behavior . . .

In other words, where psychoanalysis is relevant and succeeds with regard to explaining horror (or anything else), its successes are likely to be rare. Where psychoanalysis does appear to explain features of the experience of art horror, it is unlikely to be uniquely insightful. That is, it is likely to be based on observations available from other psychological perspectives. Psychoanalysis is largely otiose.

A problem with Carroll's discussion here is his reduction of psychoanalytic explanation to the observation of recurring patterns of behaviour. As we have argued earlier, psychoanalysis is also an orectic mode of explanation. Its explanatory power rests, in part, on the idea that our pleasures are often based on desires that are hidden from us or disguised for us. Carroll's attempt to direct the attention of film scholars away from engaging with psychoanalysis is, in the end, based on a general theoretical hostility to psychoanalysis. (p. 527) It is not based on any well-developed insights into the nature of film, the moving image, cinematic spectatorship, and the like.

Conclusion

Psychoanalytic approaches to film have focused on genres such as horror, along with feminist critiques of aspects of spectatorship such as voyeurism and misogyny that seem to cry out for psychological explanation. Psychoanalysis also played an important role in the emergence of grand theories of film in the 1970s, theories such as 'apparatus theory'. We have argued that the most significant contribution of psychoanalysis to the understanding of film is not made through such grand theories of film, but through the explanation of certain pervasive and striking features of film spectatorship, including the spectatorship of individual films and typical examples of genre films.

References

Allen, R. (1999). 'Psychoanalytic film theory'. In T. Miller and R. Stam (eds), *A Companion to Film Theory* (pp. 123–145). Oxford: Blackwell.

Allen, R. (2009). 'Psychoanalysis'. In P. Livingstone and C. Plantinga (eds), *The Routledge Companion to Philosophy and Film* (pp. 445–456). London: Routledge.

Baudry, J. L. (1986 [1970]). 'Ideological effects of the basic cinematographic apparatus'. In P. Rosen (ed.), *Narrative, Apparatus, Ideology*. New York: Columbia University Press.

Bordwell, D. (1985). *Narration in the Fiction Film*. Madison, WI: University of Wisconsin Press.

Bordwell, D. and Carroll, N. (eds) (1996). *Post-Theory: Reconstructing Film Studies*. Madison, WI: University of Wisconsin Press.

Buscombe, E., Gledhill, C., Lovell, A., and Williams, C. (1975). 'Statement: psychoanalysis and film'. *Screen 16*(4): 119–130. doi:10.1093/screen/16.4.119.

Carroll, N. (1988). *Mystifying Movies: Fads and Fallacies in Contemporary Film Theory*. New York: Columbia University Press.

Carroll, N. (1990). *The Philosophy of Horror: or Paradoxes of the Heart*. New York/London: Routledge.

Carroll, N. (1996). *Theorizing the Moving Image*. Cambridge: Cambridge University Press.

Carroll, N. (1999). 'Art, narrative, and emotion'. In M. Hjort and S. Laver (eds), *Emotion and the Arts*. New York: Oxford University Press.

Carroll, N. (2004a). 'The power of movies'. In P. Lamarque and S. Olsen (eds), *Aesthetics and the Philosophy of Art: The Analytic Tradition* (pp. 485–497). Oxford: Blackwell.

Carroll, N. (2004b). 'Afterword: psychoanalysis and the horror film'. In S. Schneider and N. Carroll (eds), *The Horror Film and Psychoanalysis: Freud's Worst Nightmares* (pp. 257–270). Cambridge: Cambridge University Press.

Clark, P. and Wright, C. (eds) (1988). *Mind, Psychoanalysis and Science*. Oxford: Blackwell.

Clover, C. (1992). *Men, Women, and Chainsaws: Gender in the Modern Horror Film*. Princeton, NJ: Princeton University Press. (p. 528)

Colby, K. M. and Stroller, R. J. (1988). *Cognitive Science and Psychoanalysis*. Hillsdale, NJ: Lawrence Erlbaum.

Copjec, J. (1989). 'The orthopsychic subject: film theory and the reception of Lacan'. *October 49*: 53–71.

Crane, J. (1994). *Terror and Everyday Life: Singular Moments in the History of the Horror Film*. London: Sage.

Creed, B. (1993). *The Monstrous-Feminine: Film, Feminism, and Psychoanalysis*. London/ New York: Routledge.

Currie, G. (1995). *Image and Mind: Film, Philosophy, and Cognitive Science*. Cambridge: Cambridge University Press.

Freud, S. (1912). 'On the universal tendency towards debasement in the sphere of love. (Contributions to the psychology of love)'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 11 (pp. 177–190). London: Hogarth Press.

Freud, S. (1916–17). *Introductory Lectures on Psychoanalysis*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vols 16 and 17. London: Hogarth Press.

Freud, S. (1919a). ' "A child is being beaten": a contribution to the study of the origin of sexual perversion'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 17 (pp. 175–204). London: Hogarth Press.

Freud, S. (1919b). 'The uncanny'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 17 (pp. 217–256). London: Hogarth Press.

Freud, S. (1927). 'Fetishism'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 21 (pp. 147–157). London: Hogarth Press.

Freud, S. (1997). *Writings on Art and Literature*, eds W. Hamacher and D. Wellbery. Stanford, CA: Stanford University Press.

Gabbard, G. (ed.) (2001). 'Introduction'. *Psychoanalysis & Film, International Journal of Psychoanalysis* (pp. 1–16). Key Papers Series, series ed. Paul Williams. London/New York: Karnac.

Gabbard, G. and Gabbard, K. (1987). *Psychiatry and the Cinema*. Chicago, IL: University of Chicago Press.

Gabbard, G. and Gabbard, K. (1999). *Psychiatry and the Cinema*, 2nd edn. Washington, DC: American Psychiatric Press.

Hopkins, J. (1982). 'Introduction: philosophy and psychoanalysis'. In R. Wollheim and J. Hopkins (eds), *Philosophical Essays on Freud* (pp. vii–xlv). Cambridge: Cambridge University Press.

Hopkins, J. (1988). 'Epistemology and depth psychology: critical notes on *The Foundations of Psychoanalysis*'. In C. Wright and P. Clark (eds), *Mind, Psychoanalysis and Science* (pp. 33–60). Oxford: Blackwell.

Humbert, D. (2013). 'Re-visiting the double: A Girardian reading of Alfred Hitchcock's *Rope* and *Strangers on a Train*'. *Contagion: Journal of Violence, Mimesis, and Culture 20*(1): 253–262.

Kemp, G. and Mras, G. (eds) (2016). *Wollheim, Wittgenstein, and Pictorial Representation: Seeing-as and Seeing-in*. London/New York: Routledge.

Lacan, J. (1977). 'The mirror stage as formative of the function of the I as revealed in psychoanalytic experience'. *Ecrits: A Selection*, trans. A. Sheridan (pp. 1–23). London: Tavistock.

Levine, M. (2004a). 'A fun night out: horror and other pleasures of the cinema'. In S. Schneider (ed.), *The Horror Film and Psychoanalysis: Freud's Worst Nightmares* (pp. 35–54). Cambridge: Cambridge University Press. (p. 529)

Levine, M. (2004). *Racism in Mind*, eds M. Levine and Tamas Pataki. Ithaca, NY: Cornell University Press.

Levine, M. (2015). 'Freud's aesthetics: artists, art and psychoanalysis'. In S. Boag (ed.), *Psychoanalysis and Philosophy of Mind* (pp. 137–162). London: Karnac.

Levine, M. (2016). 'Wollheim's ekphrastic aesthetics: emotion and its relation to art'. In G. Kemp and G. Mras (eds), *Wollheim, Wittgenstein, and Pictorial Representation: Seeing-as and Seeing-in* (pp. 241–267). London/New York: Routledge.

Levine, M. and Taylor, W. (2012). 'War as catastrophe: Jacques Callot's "Miseries of War" as moral meditation'. *Journal of Art Theory and Practice 13*(2): 157–184.

Levine, M. and Taylor, W. (2013). 'The upside of down: disaster and the imagination 50 years on'. *M/C Journal: A Journal of Media and Culture 16*:1 *Catastrophe issue*, **http://journal.media-culture.org.au/index.php/mcjournal/author/submission/586**.

McGinn, C. (2005). *The Power of Movies*. New York: Pantheon.

Metz, C. (1982). *The Imaginary Signifier: Psychoanalysis and Cinema*, trans. C. Britton, A. Williams, B. Brewster, and A. Guzetti. Bloomington, IN: Indiana University Press.

Mulvey, L. (1989). *Visual and Other Pleasures*, Bloomington, IN: Indiana University Press.

Pataki, Tamas (2014). *Wish-Fulfilment in Philosophy and Psychoanalysis: The Tyranny of Desire*. London/New York: Routledge.

Plantinga, C. (2009). *Moving Viewers: American Film and the Spectator's Experience*. Berkeley, CA: University of California Press.

Prince, S. (1996). 'Psychoanalytic film theory and the problem of the missing spectator'. In D. Bordwell and N. Carroll (eds), *Post-Theory: Reconstructing Film Studies* (pp. 71–86). Madison, WI: University of Wisconsin Press.

Radford, C. (1977). 'Tears and fiction'. *Philosophy 52*: 208–213.

Ray, R. B. (1985). *A Certain Tendency of the Hollywood Cinema, 1930–1980*. Princeton, NJ: Princeton University Press.

Rodowick, D. N. (1991). *The Difficulty of Difference: Psychoanalysis, Sexual Difference and Film Theory*. New York: Routledge.

Schneider, S. J. (1997). 'Uncanny realism and the decline of the modern horror film'. *Paradoxa: Studies in World Literary Genres 3*(3/4): 417–428.

Schneider, S. J. (2000). ' "Monsters as (uncanny) metaphors": Freud, Lakoff, and the representation on monstrosity in cinematic horror'. In A. Silver and J. Ursini (eds), *Horror Film Reader* (pp. 167–191). New York: Limelight Editions.

Schneider, S. J. (ed.) (2004a). *The Horror Film and Psychoanalysis: Freud's Worst Nightmares*. Cambridge: Cambridge University Press.

Schneider, S. J. (2004). 'Introduction: psychoanalysis in/and/of the horror film'. In S. J. Schneider (ed.), *The Horror Film and Psychoanalysis: Freud's Worst Nightmares* (pp. 1–15). Cambridge: Cambridge University Press.

Sklarew, B. (1999). 'Freud and film: encounters in the *weltgeist*'. *Journal of the American Pyschoanalytic Association 47*: 1239–1247.

Smith, M. (1995). *Engaging Characters: Fiction, Emotion and the Cinema*. Oxford: Oxford University Press.

Sontag, S. (1965 [1979]). 'The imagination of disaster'. In *Against Interpretation and other Essays* (pp. 209–225). New York: Dell.

Sterba, R. (1940). 'The problem of art in Freud's writings'. *Psychoanalytic Quarterly 9*: 256–268.

Studlar, G. (1985). 'Masochism and the perverse pleasures of the cinema'. In B. Nichols (ed.), *Movies and Methods*, vol. 2. Berkeley, CA: University of California Press. (p. 530)

Studlar, G. L. (1988). *In the Realm of Pleasure: Von Sternberg, Dietrich, and the Masochistic Aesthetic*. Urbana, IL: University of Illinois Press.

Tudor, A. (1997). 'Why horror? The peculiar pleasures of a popular genre'. *Cultural Studies 11*: 443–463.

Wolfenstein, M. and Leites, N. (1950). *Movies: A Psychological Study*. Glencoe, IL: Free Press.

Wood, R. (2004). 'Foreword: "What lies beneath?" ' In Schneider (ed.), *The Horror Film and Psychoanalysis: Freud's Worst Nightmares* (pp. xiii–xviii). Cambridge: Cambridge University Press.

Žižek, S. (1992). *Enjoy Your Symptom! Jacques Lacan in and out of Hollywood*. New York: Routledge.

Notes:

([1]) Freud, however, was not an enthusiast. Gabbard (2001: 1–2) claims that 'The cinema and psychoanalysis have a natural affinity.' But he goes on to say that 'The marriage between movies and psychoanalysis occurred in *spite of* Sigmund Freud. As far as we know, Freud had little regard for the cinema as an art form and appeared almost oblivious to the development of movies during his lifetime (Sklarew 1999). His attitude was perhaps best illustrated when Hollywood producer Samuel Goldwyn offered him a $100,000 fee to consult on a film he was planning to shoot in 1925. Freud rejected the offer without a second thought. The *New York Times* of 24 January 1925 displayed the following headline: Freud Rebuffs Goldwin: Viennese Psychoanalyst Is Not Interested in Motion Picture Offer (Sklarew 1999: 1244).'

([2]) Tudor (1997: 444) distinguishes between the following two questions. ' "What is it about people who like horror?" and "what is it about horror that people like?" ' Psychoanalytically speaking, however, these are just two sides of the same coin.

([3]) Levine (2015).

([4]) Metz (1982); Mulvey (1989). Other examples include: Baudry (1986), Clover (1992), Copjec (1989), Creed (1993), Rodowick (1991), Schneider (2000), Studlar (1988), and Žižek (1992).

([5]) The publication of David Bordwell's *Narration in the Fiction Film* (1985) is the founding moment of the cognitivist turn in film theory. The pervasiveness of cognitivist criticism of psychoanalytic approaches to film is on display in such places as the *Routledge Companion to Philosophy and Film* (2009). The entry on psychoanalysis in this work, by Richard Allen, is robustly critical and sometimes dismissive. We discuss Allen's views later.

([6]) Prominent cognitivists in film studies include Bordwell (1985), Carroll (1996), Currie (1995), Plantinga (2009), and Smith (1995).

([7]) See Damien Freeman and Elisa Galgut, both in this volume. Also see Kemp and Mras (2016) and Levine (2016).

([8]) See Sterba (1940: 257): 'The … complemental relationship between the two sources of information consists then in the theoretical enlightenment about art and the artist to be found in papers devoted to the psychoanalytic theory of neurosis, while in those papers devoted to an examination of specific objects of art and artists we are given insight into more general analytic theories and methods.'

([9]) See Pataki's (2014) account of wish-fulfilment and desire in philosophy and psychoanalysis, and chapter 17, this volume.

(^{10}) See Buscombe et al. (1975: 119, 125–6) for criticism of psychoanalytic approaches to film published in *Screen*, the leading academic journal for the study of film and television. 'Our reservations are in three main areas: the unproblematic acceptance of psychoanalysis implicit in the way it has been presented in *Screen*; the intelligibility of the various expositions and applications of it; and the validity of the attempts made to apply it directly to the cinema' (119). The objections appear to amount to rejection of the validity of psychoanalysis in its entirety.

(^{11}) Gabbard (2001: 8) says 'Part of the appeal of the horror and science fiction genres is related to the audience's vicarious mastery of infantile anxieties associated with earlier developmental crises. The audience can re-encounter terrifying moments involving early anxieties while keeping a safe distance from them and knowing that they can survive them.'

(^{12}) Sontag discusses many films. They include *War of the Worlds* (1953), *Conquest of Space* (1955), *This Island Earth* (1955), *The Mysterians* (1957), *The Day the Earth Caught Fire* (1962).

(^{13}) It is a common mistake to think this ahistoricity excludes a sensitivity to context. Schneider (2004: 11–12) takes up this issue of historicity in discussing Crane (1994) and Tudor (1997). He writes: 'Jonathan Lake Crane … charges those theorists who offer depth-psychological explanations of canonical horror film monsters and narratives … with ahistoricity: "In irrevocably linking horror to the unconscious we dismiss, all too hastily, the possibility that horror films have something to say about popular epistemology, about the status of contemporary community, or about the fearsome power of modern technology" (Crane 1994: 29).' But the psychoanalytic linking of horror to the unconscious entails no such dismissal. On the contrary, psychoanalysis supports all of Crane's claims with a variety of explanatory theses. Similarly, Schneider (2004a: 12) notes that 'Tudor finds fault with those "universalizing" explanations of the pleasure viewers get from watching horror films on the grounds that "it is only possible to speak of the appeal of a genre in a particular sociotemporal context … [P]sychoanalytic models, arguably already reductive, will be particularly misleading, conceptually inclined to neglect the variability of audience responses in the name of a spurious generality"(1997: 456).' This misunderstanding was addressed earlier.

(^{14}) See Levine and Taylor (2012, 2013).

(^{15}) See Levine (2004a) for a response to criticisms of psychoanalytic film theory's approach to horror.

(^{16}) Carroll (1990).

(^{17}) See Schneider (1997, 2000) for an explanation of how (i) the 'return of the repressed', and (ii) the reconfirmation of previously surmounted infantile beliefs addresses the question of the attraction of horror. Schneider also explains—again in terms of the above

—why certain kinds of monsters and horror films lose their appeal and fail to frighten by failing to evoke the 'uncanny'.

([18]) A doppelganger ('double walker') is a double or second self. In literature and film it is often portrayed as a kind of copy or mirror image of the protagonist and may symbolically represent an archetype. See David Humbert (2013) for development of the idea of the significance of doubling in Hitchcock's work.

([19]) It should be noted that what Carroll calls the 'hermeneutical approach' is not what is generally meant by a hermeneutical approach (e.g. compare it with Paul Ricœur's account), but is instead a simple form of contextualism. Psychoanalysis can be and is used hermeneutically, but not generally in the ways that Carroll claims it can be appropriately used. The ways Carroll thinks appropriate are indifferent to the truths of psychoanalysis. They are based only on explicit and implicit authorial intent, but not intent as psychoanalytically interpreted.

([20]) Schneider (2004: 11) puts the point in this way: 'Just as "one should not assume, *prima facie*, that either Freudian psychology or one singular version of psychoanalytic theory is the key to understanding a text" (Allen 1999: 142), one should not assume that any version of psychoanalysis, not even any *combination* of versions, is the key to understanding *every* horror text. Most critics object to what they perceive as psychoanalytic horror film theory's unsupportable claims of explanatory sufficiency. This objection appears in a number of forms and cuts across the distinction I have been drawing between a minimal, epistemically neutral use of psychoanalytic theory and a use of such theory which depends on the truth of at least some of the particular version's substantive theses'.

([21]) See Hopkins (1988) and Levine (2004a: 35–40) for a critique of the claim that psychoanalysis is unfalsifiable.

([22]) Stephen Prince (1996: 72–3) states that 'the primary and to my mind insurmountable problem with basing general theories of spectatorship on psychoanalysis is that such theories must remain unsupported because psychoanalysis is a discipline without reliable data'. Prince endorses Colby and Stroller (1988: 3, 29), who claim that 'psychoanalytic evidence is hearsay, first when the patient reports his or her version of an experience and second when the analyst reports it to an audience … Reports on clinical findings are mixtures of facts, fabulations, and fictives so intermingled that one cannot tell where one begins and the other leaves off'. See Levine (2004a) for a critique of these claims.

([23]) Tudor (1997: 450) cites *The Exorcist, Carrie, Alien, The Brood*, and *The Hunger* as examples of films in which the monstrous-feminine 'does play an important role'. It is unclear why he makes this claim given his apparent rejection of feminist film theory that employs 'structural psychoanalysis'.

([24]) See Hopkins's (1982: xxi) discussion of the table cloth lady.

Psychoanalysis and Film

Damian Cox

Faculty of Society and Design, Bond University, Australia

Michael Levine

The University of Western Australia, Australia

Oxford Handbooks Online

Abstract and Keywords

This introduction provides an overview of the chapters in this section, which explores some of the important contributions of psychoanalysis to our understanding of religion, with particular emphasis on Sigmund Freud's views. In The Future of an Illusion (1927), Freud argues that religion—or more specifically, beliefs in the Judaeo-Christian tradition and its prehistory—is an 'illusion', an idea that is not necessarily false, but one that is produced by the wish for it to be true. Each of the chapters agrees with the notion that there is a close connection between religious belief and desire, and addresses Freud's account of the origin of religion in structures of subjectivity. Topics include Jacques Lacan's theory of religion, the implications of psychoanalytic theories of subjectivity for philosophy of religion, the epistemology of religious belief in relation to the epistemology of psychoanalysis itself, and Freud's supposed view that religion expresses a 'historical truth'.

Keywords: psychoanalysis, religion, Sigmund Freud, religious belief, desire, subjectivity, Jacques Lacan, wish, philosophy of religion

Michael Lacewing and Richard G. T. Gipps

THE standard interpretation of Freud's account of religion goes something like this. In The Future of an Illusion (1927), Freud compares religion—or more specifically, beliefs in the Judeo-Christian tradition and its prehistory—to an 'illusion', an idea that is not necessarily false, but one that is produced by the wish for it to be true. Let us think about the human situation independent of the claims of religion. What would human beings wish the universe to be like, given our situation? The answer, Freud argues, is just how religion describes the world as being. In other words, religion presents an understanding of the universe that we can understand as a wish fulfilled.

Religion: Introduction

The origins of the wish fulfilled by religious belief are two-fold. First, in the history of humanity, we find ourselves part of an overpowering, uncaring nature. Thus, we need ways to propitiate the uncontrollable forces we face. And so we seek to be able to control the forces, e.g. through influencing a god that can do this—the god of the sea or wind or a God who is God of all things. Or, we can seek to align ourselves with God, to live a good life that will protect us from the worst effects of these forces. Our security rests in God's love for us—if it does not save us physically, it does so in a deeper sense. Second, in our individual history, as children, we each experience our dependency, our inability to take care of ourselves. Our wishes are met by our parents, but we soon realize their fallibility and limitations. Hence, we arrive at the idea of a much more powerful and perfect parent in God. Thus, argues Freud, we have a naturalistic, psychological explanation of religion—we can explain the origin of its claims in our wishes and the origin of our wishes in our experience of reality.

However, we can turn Freud's approach on its head. Suppose God does exist and created us. What would human beings wish for—or better—what would human beings desire deeply? God—the fulfilment of our nature would be to stand in good relationship with God. Here, however, the desire is realistic, just as the child's desire to be looked after by its parents is realistic, necessary for its life and flourishing. Even if Freud (p. 534) shows that religious belief originates, at least in part, in our wishes, he does not show that those wishes, and so religious belief, originate simply in response to our helplessness in a difficult and dangerous world.

If we accept this much—that religious belief is closely connected to desire—as each of the chapters included here does, there are at least these three paths of thought leading on from here. First, we may seek to develop or challenge Freud's account of the origin of religion in structures of subjectivity. Each of the three chapters included here takes us down different branches of this path that reclaims religion from Freud's negative evaluation of it (see also Black 2006; Hinshelwood 1999). Both Cottingham and Blass, in different ways, review and deepen the standard interpretation of Freud before going on to discuss, respectively, Jung and the resources of modern cognitive psychology and neuroscience, and recent psychoanalytic theories in the British and American traditions that regard religious belief positively, while Boothby provides us with an overview of Lacan's theory of religion.

Second, we can develop the question of whether these accounts of religious belief undermine, support, or simply bracket as unimportant the truth of religious belief, or again whether they have some effect on our understanding of what it could be for religious belief to be 'true'. This is the main focus of Blass's discussion.

Third, we may ask what further implications—beyond adding possible reasons for or against belief in the existence of God—psychoanalytic theories of subjectivity have for philosophy of religion. There are, of course, different philosophies of religion, drawing on different traditions in philosophy. Boothby notes the development of ideas from Kant and Hegel on religion and subjectivity in the thought of Lacan and the modern Lacanian

Žižek. An important project in Cottingham's previous work (2003, 2014), of which he provides a clear synopsis here, has been to consider the implications for philosophy of religion in the analytic tradition (see also Clack 2008). In particular, he criticizes the dominant 'epistemology of control' that reflects only the conscious, detached, analytic way that we may respond to questions about God. Our best psychological theories indicate that this needs integrating with an 'epistemology of receptivity' that reflects the direct, intuitive, imaginative, and typically unconscious ways in which we constantly engage with the world through all our cognition, and may be of yet greater importance in acquiring the relevant evidence for thinking meaningfully about religion. Here, notes Cottingham, the epistemology of religious belief (and its necessary reflection in philosophy of religion) mirrors the epistemology of psychoanalysis itself.

This is a conclusion that Blass shares. After carefully demonstrating that Freud's later views on religion changed from those expressed in *The Future of an Illusion*, she criticizes recent theories for failing to take the truth of religious claims seriously, yet it was exactly this point that vexed Freud, and his search for truth is not something we should lightly set aside (see Gipps and Lacewing, 'Introduction: Know Thyself' in this volume). Blass shows that by Moses and Monotheism (1939), Freud arrives at the view that religion expresses an 'historical truth', an impression made on the mind by past reality, but which does not accurately reflect that reality, because the mind cannot do so at (p. 535) that time as the events are too primal. The origin of religion in the history of humanity rests in such events, Freud thinks, but this new account of mental functioning has implications for primal events in each individual's life as well. A sense of conviction in the 'rightness'—or truth—of an idea cannot be dismissed as wishful, but may reflect some fundamental aspect of human nature and experience—and this is something that the psychoanalysis must be attentive to more generally. It is a surprise to find this thought in Freud's late work, as it is more often presented by Freud's critics as an objection to the 'standard story'.

Lacan's thought on religion is no less full of surprises, as richly brought out by Boothby. Lacan understands the mind and its operations along three axes or registers. To oversimplify hugely, 'the imaginary' structures the ego, its initial division as 'me' (*moi*) from the true 'subject' (*je*) of desire, and its operations of defence and transference. 'The symbolic' concerns the sphere of language and rules/laws more generally, generated through and regulating our encounter with 'Others'. 'The real' involves our encounter with what is incomprehensible, both outside and inside us. Where does God (as thought by us) fit in? Is God part of the incomprehensible real, as Kant argued? Or is God the ultimate 'Other', providing regulation for our lives? Boothby argues that both answers are correct, unified and coordinated by Lacan's concept of 'the Thing', that which is both beyond representation and yet can only be thought—symbolized—at all as precisely being beyond representation. In other words, it is symbolized as a question mark or an 'X'. The real and the symbolic thus arise together in human subjectivity, with the first incomprehensible real being the desire of the Other. Our relation to the real and the Other is ultimately one of desire, not—as Kant held—of thought. Along the way, Lacan

argues that this account not only explains religion but specific practices, such as sacrifice, and Christian doctrines, such as the Incarnation.

The important contributions of psychoanalysis to our understanding of religion are not only, then, Freud's negative critique, but bringing to our attention the deeper roots of religious belief in human subjectivity. Philosophy of religion can no longer rest content with examining religion in relation to rationality but must come to grips with unconscious imagination, emotional response, and desire. It is interesting to note the steps already being taken in this direction (e.g. Coakley 2013; Ellis 2018).

References

Black, D. M. (ed.) (2006). *Psychoanalysis and Religion in the 21st Century: Competitors or Collaborators?* London: Routledge.

Clack, B. (2008). 'After Freud: phantasy and imagination in the philosophy of religion'. *Philosophy Compass 3*: 203–221.

Coakley, S. (2013). *God, Sexuality and the Self*. Cambridge: Cambridge University Press.

Cottingham, J. (2003). *The Spiritual Dimension*. Cambridge: Cambridge University Press.

Cottingham, J. (2014). *Philosophy of Religion: Towards a More Humane Approach*. Cambridge: Cambridge University Press. (p. 536)

Ellis, F. (2018). 'Philosophy of religion and desire'. In L. Hickman and R. Manning (eds), *Philosophy of Religion: Whence and Whither?* London: Bloomsbury.

Freud, S. (1927). 'The future of an illusion'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 21 (pp. 1–56). London: Hogarth Press.

Freud, S. (1939). 'Moses and monotheism: Three essays'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 23 (pp. 1–137). London: Hogarth Press.

Hinshelwood, R. (1999). 'Introduction'. In S. M. Stein (ed.), *Beyond Belief: Psychotherapy and Religion*. London: Karnac Books.

Michael Lacewing

Dr Michael Lacewing is a former Vice-Principal Academic and Reader in Philosophy at Heythrop College, London, an Honorary Reader in Research Department of Clinical, Educational and Health Psychology at University College, London, and a teacher of philosophy and theology at Christ's Hospital School, Sussex. He has published widely in philosophy of psychoanalysis, metaethics and moral psychology,

alongside writing textbooks for A level philosophy and training in Philosophy for Children (P4C).

Richard G.T. Gipps

Dr Richard Gipps is a clinical psychologist in private psychotherapy practice in Oxford, UK, and an associate of the Philosophy Faculty at the University of Oxford. He convenes the Philosophy Special Interest Group of the Institute of Psychoanalysis, the Oxford Interdisciplinary Seminars in Psychoanalysis, and the Making the Unconscious Conscious seminar series. His research interests lie in psychoanalysis, psychosis, existential phenomenology, and Wittgenstein.

Subscriber: UC - Los Angeles (UCLA); date: 17 March 2019

Oxford Handbooks Online

Psychoanalysis and Religion 🔓

John Cottingham

The Oxford Handbook of Philosophy and Psychoanalysis
Edited by Richard G.T. Gipps and Michael Lacewing

Abstract and Keywords

Acknowledging the layers of the mind below the level of overt consciousness can lead to very divergent accounts of religious belief. One response—taken by Freud himself—argues that religious belief should be abandoned as unavoidably contaminated by unconscious motivations (e.g. an infantile longing for security) that distort our rational judgement. By contrast, Jung maintains that religious thinking is shaped by unconscious structures (the 'archetypes'), which can play a vital role in the development of an integrated human personality. This chapter examines these contrasting psychoanalytic interpretations of religion, and then explores more recent accounts of the workings of the human psyche and how they affect the status of religious belief. A concluding section discusses some general implications of all this for the epistemology of religious belief and the way in which philosophy of religion should be conducted.

Keywords: Freud, Jung, unconscious, illusion, archetypes, Graham Ward, Iain McGilchrist, brain hemispheres, receptivity, detachment

Introduction

A large part of the theory and practice of psychoanalysis, from Sigmund Freud onwards, has been concerned with the treatment of disturbed and troubled patients seeking help with their symptoms, but the philosophical importance of the Freudian revolution is very far from confined to the domain of the pathological or the neurotic. In effect what Freud challenged was a simplistic but widely held model of the mind as a kind of transparent goldfish bowl inside which straightforwardly identifiable beliefs and desires float around, ready to inform our actions and choices. In the wake of Freud, it has become much harder to be confident about the image of ourselves as self-sufficient and autonomous rational agents whose decisions are based solely on what is straightforwardly accessible to the reflective mind. This does not mean that the Freudian revolution in our conception of ourselves undermines the very possibility of rational thought—if that were the case, psychoanalytic thought would be self-refuting, since it would undermine the possibility of its coherent articulation. What *is* entailed is that we should give up the naïve conception of our mental powers and capacities as transparent tools of reason, and start working towards a more nuanced conception, according to which uncovering the truth about ourselves and our relation to the world must be approached in a spirit of humility and receptivity that acknowledges the intensely complex and problematic nature of the instrument which we must use to undertake that task—the human mind.

These general lessons of the Freudian revolution evidently have application to the domain of religious belief, along with many other areas of human thought. But acknowledging the layers of the mind that operate below the level of overt consciousness can lead to very divergent accounts of the status and validity of religious beliefs and attitudes. One way to go—the route that Freud himself took—is to argue that religious belief should be abandoned, insofar as it is unavoidably contaminated by unconscious drives and motivations (an infantile longing for security, for example) that distort our rational judgement. A quite opposite approach, exemplified by that of Freud's onetime disciple Carl Jung, is to maintain that religious thinking is typically shaped by unconscious forms and structures (what Jung called the 'archetypes') which, so far from being generators of neurosis, can play a vital role in the development of a healthy and integrated human personality. We shall look in more detail at these two influential but strongly contrasting psychoanalytic interpretations of religion in the next two sections, before going on, in the fourth section, to explore more recent accounts of the workings of the human psyche and how they may affect the status of religious belief. The fifth and final section will aim to tease out some general conclusions about the relationship between psychoanalysis and religion, and the implications of this for the epistemic status of religious belief, and the way in which the philosophy of religion should be conducted.

Freud's Critique of Religion

Freud's attack on religion begins by drawing attention to our human helplessness before the 'majestic, cruel and inexorable powers of nature' (Freud 1927: 195). These powers include both external forces (earthquakes, floods, hurricanes) and the equally threatening internal forces (lust, anger, brutality) arising from our own nature. Freud sees religion as an attempt to mitigate our defencelessness by endeavouring to 'adjure, appease, bribe' or otherwise influence a celestial father figure, who will protect us from suffering, and impose justice on a seemingly chaotic and terrifying universe (1927: 196).

The vulnerability of the human condition, and the fact that since time immemorial humans beings in extremis have resorted to a variety of supposed divine powers and forces to rescue them when all else fails, are familiar enough themes which have been commented on by many writers, including Jean-Paul Sartre:

> When the paths traced out become too difficult, or when we see no path, we can no longer live in so urgent and difficult a world. All the ways are barred. However, we must act. So we try to change the world, that is, to live as if the connection between things and their potentialities were not ruled by deterministic procedures, but by magic
>
> (Sartre 1975: 58–9).[1]

The same general line had been taken much earlier by another stern critic of religion, David Hume. What prompts us to suppose there is a God, according to Hume, are 'the ordinary affections of human life' such as the 'dread of future misery' and the 'terror of death' (Hume 1757: section 2).

But to explain the religious impulse simply in terms of human vulnerability and helplessness leaves something out: no doubt people earnestly desire to be rescued when in trouble, but that in itself does not seem to account for the strength and pervasiveness of the human belief in the divine. When in grave distress we might *like* to deceive ourselves into thinking the world is determined not by natural forces but by magic (as Sartre phrases it), but this does not in itself explain how widespread and successful such self-deception (if it is indeed that) has become. Here Freud contributes something crucially important by addressing himself to the psychological question of *how* the belief becomes so powerfully entrenched in the minds of so many religious adherents. The human psyche, he argues in *The Future of an Illusion*, is *already predisposed*, as a result of the traces left by our forgotten experience as infants, to conjure up the image of a powerful protector to rescue us from our helplessness.[2] For when we encounter threats and dangers:

… this situation is nothing new. It has an infantile prototype of which this is in fact only the continuation. For once before one has found oneself in a similar state of helplessness: as a small child, in relation to one's parent. One had reason to fear them, and especially one's father; and yet one was sure of his protection against the dangers one knew. Thus it was natural to assimilate the two situations. Here, too, wishing plays its part, as it does in dream-life …

(Freud 1927: 196).

So just as with the strange deliverances of dreams, what is planted in our consciousness has a resonance, a power that takes hold of us quite independently of the normal criteria of reasonable evidence and rational judgement. The mind is in the grip of an *illusion*, but this is not just a mistake, or a piece of deliberate self-deception. Rather, layers of mentation working beneath the level of explicit awareness or rational reflection have been activated by our helplessness in the face of the perils we face as adults, and we revert, without being consciously aware of what is going on, to the infantile state of fear and dependency which is ineradicably linked to the yearning for security and the hope of parental protection. Only as a result of delving into the deeper workings of the mind right back from early childhood does the full explanation of the process come to light. This is the background that enables Freud to declare with such confidence 'the derivation of religious needs from the infant's helplessness and the longing for the father aroused by it seems to me incontrovertible' (Freud 1929: 260). So we arrive at the famous Freudian diagnosis: religion is an illusion born of helplessness and fear.

Freud makes it clear, however, that he does not intend his psychoanalytic diagnosis of our longing for protection to be a logical demonstration of the falsity of the religious world view. That would be to commit the 'genetic fallacy' (the logical fallacy of confusing the causal origins of a belief with its justification or lack thereof). Illusions, as Freud concedes, are not *necessarily* erroneous: 'A middle-class girl may have an illusion that a prince will come and marry her … and a few such cases have actually occurred' (Freud 1927: 213). But Freud argues that it is characteristic of illusions in his sense that they are held on to without regard for rational justification; further, they characteristically stem from (indeed are generated by) the wishes or needs of the believer. So it is a short step from this to the conclusion that Freud is aiming at: religion is an infantile piece of wishful thinking that we need to grow out of.

Yet on further reflection the implications of Freud's critique are by no means as damaging as might at first appear. The believer might well concede to Freud that our infantile helplessness leaves a lasting stamp on the psyche, but go on to insist that this can scarcely be the whole story. For beyond any mere desire for protection (Freud's 'longing' for the father figure), it seems hard to deny that the religious impulse is in large part connected with the powerful yearning human beings have for meaning and purpose in their lives. Now it could be proposed, as is done by many secularists, that meaning and purpose must be found in the chosen activities and pursuits—intellectual, artistic, social, familial, and so on—which are the components of a worthwhile human life. But, without

denying the value and meaningfulness of such activities and pursuits, it may be argued that they cannot in themselves bear all the weight of satisfying our human hunger for meaningfulness. One way of putting this is to say that to be human is to have a characteristic restless, a sense of incompleteness, such that even were all our specific needs and goals to be satisfied (for food, for shelter, for company, for recreation, for satisfying relationships, for creative activities, and so on), there would always remain a longing for something more—something that will provide an 'ultimate grounding' for our lives, or give us a sense of 'ontological rootedness' (May 2011: 7).

God, for the religious believer, is the ultimate source of being and value towards which we yearn, and which alone can satisfy the existential longing which is part of the nature of dependent and contingent beings such as us.[3] Pointing this out does not of course vindicate belief in God, nor does it of itself refute deflationary Freudian-style explanations of it, but it may at least open up the possibility that religious belief connects with something in our human nature of deeper significance than a mere neurotic or infantile impulse. Certainly there are many places in Scripture where the strange open-ended longing of the human spirit is underlined ('Like as the hart desireth the water-brooks, so longeth my soul after Thee, O God' (Psalm 42 [41]: 1)); and the theme is reiterated in seminal Christian writers such as St Augustine ('You have made us for Yourself, and our heart is restless until it finds repose in You', and Dante ('In his will is our peace')).[4] The thought in such passages is not merely that religious devotion provides peacefulness of mind, in the sense of securing some kind of tranquillizing or calming effect; rather, the idea is that God is the source of genuine value, and that orienting ourselves towards that source bestows ultimate meaning on our human existence and enables us to find true fulfilment even in the face of danger and turmoil. Augustine and Dante acknowledge our vulnerability, but manage to construe it as a corollary of our creatureliness, so they can end up celebrating it as a cause for joyful affirmation of our creator. Freud by contrast sees our vulnerability as a condition which scares us so much that we desperately fantasize that we have found a way of assuaging it—even though in fact the power we appeal to has no reality outside the human psyche. But as to which of the two accounts reflects the way things actually are, this remains to be determined; so however disconcerting Freud's analysis may initially be for the religious believer, it seems clear that it cannot finally settle the matter.

Carl Jung and the Importance of Symbolic Thought

While Freud's view of religious belief places it under the same general heading as the neurotic, the infantile, and the disturbed, or at any rate as falling short of the standards of balanced judgement to which we aspire as reasonable adults, Carl Jung, though concurring with Freud that such belief has roots buried deep in the human psyche, took a very different view of the resources of the unconscious mind, regarding them in a potentially much more benign light.[5] For Jung, Freud's dismissal of the religious impulse as infantile fails to recognize the imaginative and symbolic role of religious modes of thought and expression, and their possible role in the healthy development of the human personality. Crucial here is the idea of 'individuation' as Jung terms it, the 'process by which a person becomes a psychological "in-dividual", that is, a separate, indivisible unity or whole' (Jung 1939: 275). Jung sees human psychological development in terms of a struggle to achieve internal balance and psychic integration, where integrating the conscious and unconscious elements of the self is a precondition for psychic health or wholeness;[6] and religious imagery and symbolism, according to Jung, perform a vital function here (Jung 1952). The process of individuation requires modes of thought and expression that operate not just on the surface level of explicit assertion, but which carry deep imaginative resonances that are vital for our psychological balance and harmony. To give but one example of this, the figure of Christ functions for Jung as an 'archetype of the self', a deeply resonant image of the perfectly unified and integrated human being (Jung 1938; Jung 1951: 183). From this perspective, as Michael Palmer aptly puts it in his account of the Jungian position:

> Religion, far from being neurotic, is revealed as a constant and evolving process in the development of the psychic personality ... Religious symbols ... open up a psychic level ... that is primordial and ... of supreme value for the present and future development of the human psyche
>
> (Palmer 1997: 110–11).

Jung's ideas have encountered considerable philosophical opposition (as indeed have those of Freud). Many contemporary analytic philosophers are supporters of what Brian Leiter has called the 'naturalistic revolution in philosophy', according to which philosophy should 'adopt and emulate the methods of the successful sciences' (Leiter 2004: 2–3). Such philosophers often tend to be sceptical about the very idea of the unconscious mind, and *a fortiori* the Jungian idea of the archetypes, on the grounds that the theories that invoke such ideas lack the kind of hard scientific warrant demanded by today's dominant naturalistic paradigm. In response to this kind of critique, defenders of the Jungian approach have two possible lines of defence. One is to argue that there is in fact hard empirical evidence, for example from cognitive science and developmental psychology, that can be used to support the Jungian hypothesis of the role of symbols and archetypes

in psychic integration (Knox 2003). Another response would be to take issue with the doctrine that the methods of science are the only valid way of uncovering the truth. Thus Thomas Nagel has argued, in the case of Freud, that irrespective of what we think about the authority of analysts or the clinical evidence for their theories, there is an 'evident usefulness of a rudimentary Freudian outlook in understanding ourselves and other people, particularly in erotic life, family dramas, and what Freud called the psychopathology of everyday life' (Freud 1901; Nagel 1994). And similarly one could argue that Jung's ideas, like those of Freud, are best assessed not as contributions to science, but in a more 'hermeneutic' way—that is, as ways of enriching our understanding of the human predicament, and the deeper significance of our thoughts and feelings and beliefs.[7], [8]

To suggest that Jung's ideas may enrich our understanding without qualifying as contributions to science is not at all to dismiss or downgrade the value and importance of scientific inquiry or scientific methods. One can be a genuine and wholehearted admirer of the achievements of *science* while at the same time resisting the false allure of *scientism*—the dogma that scientific methods give us everything we need to understand all aspects of reality. To be sure, we live in, and are an integral part of, the physical world constituted by the particles and forces studied by science—that is undeniable. But when it comes to understanding aspects of human life such as religious experience (and the same goes for poetic or artistic or moral experience, or even our ordinary human interactions with each other) we patently need other categories than those of the physical sciences; for even the fullest and most detailed print-out of the relevant particle collisions and biochemical processes will tell us nothing about the *human significance* of these processes and events.[9]

In addition to the physical sciences there are of course the social sciences (including for example economics, sociology, and psychology), and there are continuing debates as to how far such disciplines meet the standards of the 'hard' physical sciences (in matters such as experimental repeatability, verifiable prediction, mathematical modelling, and so on). There is probably no simple answer to this question, since the term 'social science' covers a large array of divergent disciplines and inquiries, whose methods manifest varying degrees of precision and rigour. A particular issue as regards psychology is the inevitable reliance on reports by individual human subjects of their thoughts, feelings, sensations, beliefs, and desires, thus making reference to what is, according to some philosophers, an irreducible domain of qualitative subjective experience that resists subsumption or explanation in objective scientific terms (Nagel 1974). But however that may be, psychoanalytic approaches to psychology present special additional problems, insofar as the 'data' being studied come to light in the context of special relationship with the analyst—an issue which led even Freud, despite his attraction to the scientific model, to remark on how far psychoanalysis diverges from normal scientific procedures. We shall return to this issue in 'Philosophizing about Religion and the Layers of the Human Psyche'.

At all events, when we come to religious feelings, beliefs, attitudes, and experiences, it is apparent that these have a characteristic depth and complexity that can seldom if ever be conveyed in a set of straightforward factual propositions laid out for our assessment and awaiting verification. For religious thoughts and ideas operate within a rich and complex web of *associations*, carrying manifold metaphorical and symbolic echoes which may often have powerful effects on us in ways that are not fully transparent to consciousness. This, as Jung sees it, is the key to the peculiar resonance and power of the images and icons that inform the thoughts and ideas of religious believers, and what explains their role in the search for integration and healing within the troubled human psyche. None of this of course means that we should uncritically accept all or any of Jung's ideas about the role of religious concepts; but at least it reminds us of the context in which his theories are meant to operate, and within which they need to be evaluated.

Whatever conclusions one finally reaches about the Jungian theory of archetypes, important questions remain about Jung's general approach to religion. The foremost among these is the objection that the Jungian approach leads to a kind of psychologizing or subjectivizing of religion, where the question of the truth or validity of any given religious outlook (Christian theism, for instance) boils down to no more than the question of whether certain archetypal images (such as that of God the Father, or Christ the Son, for example) have a transformative power within the human psyche (Palmer 1997: 187, 196). Jung's own response to this type of criticism was that his role as a psychologist was not to make pronouncements about the existence or non-existence of transcendent realities, but simply to describe the role of certain fundamental and universal images and symbols in human development:

> We know that God-images play a great role in psychology, but we cannot prove the [actual] existence of God. As a responsible scientist, I am not going to preach my personal and subjective convictions which I cannot prove … To me, personally speaking, the question whether God exists at all or not is futile. I am sufficiently convinced of the effects man has always attributed to a divine being. If I should express a belief beyond that … it would show that I am not basing my opinion on facts … I am well satisfied with the fact that I know experiences which I cannot avoid calling numinous or divine
>
> (Jung 1956).

This makes quite clear the restricted scope of Jung's position: it insists that religious concepts and images play a crucial role in the development of the human personality and its search for integration, but leaves completely open the question of whether there is some objective reality—something 'external' or independent of the subjective structure of the human psyche—to which those concepts and images refer.

To sum up our necessarily compressed and selective account of the contrasting attitudes of Freud and Jung to religion—the former's highly negative, the latter's much more positive—what emerges is that in neither case do their suggested findings about the workings of the human psyche in themselves either establish or refute the truth of the

religious outlook. The partly hidden motivation for religious belief may, if Freud is right, be an infantile one; but as we have seen, that in itself does not logically entail the falsity of such a belief. And the symbols and images drawn from the unconscious mind may, if Jung is right, exert a powerful psychological influence on the human quest for integration; but that, as just noted, still leaves open the real existence or otherwise of the God that is the object of religious belief.

In the wake of these two seminal thinkers, however, one thing at any rate should be clear: that any philosophical attempt to address the fundamental questions of religious belief will find it hard to carry conviction unless it takes some account of the complexity that lies beneath the seemingly transparent surface of propositional assent to religious claims and doctrines. No account of religious belief and experience is going to look plausible unless it acknowledges the complexity of the human mind—the strata of hidden longings and needs and the manifold symbolic forms and images resonating deep within the human psyche. To some of the more recent attempts to address that complexity we shall now turn.

The Complexity of Belief

The contemporary debate over the validity of religious belief tends to play out in a curiously abstract and rationalistic way. The implicit assumption is that the participants are detached evaluators, judiciously examining the 'God hypothesis' (as Richard Dawkins calls it (Dawkins 2006: ch. 2)), scrutinizing the supposed evidence, and weighing up the arguments for and against. Particularly among anglophone philosophers in the analytic tradition, where psychoanalytic ideas have tended to be resisted or ignored by many practitioners, the claims of religion are implicitly construed as rather like scientific claims, suitable subjects for purely intellectual disputation, the province, as it were, of scholarly discussion in the seminar room. But a number of recent writers have started to challenge this austere and bloodless picture of religious belief.

Philosophers have argued endlessly about the epistemic status of religious and other kinds of belief, and what entitles some beliefs to the accolade 'knowledge', but comparatively few have paid attention to 'what lies beneath'—to the 'archaeology of belief', as the British theologian Graham Ward has called it. Ward argues that believing or disbelieving something involves far more complex processes than the scrutiny and evaluation of factual evidence. There are much 'deeper layers of embodied engagement and reaction', where we are touched 'imaginatively, affectively and existentially' (Ward 2014: 7, 10, 31). Drawing on empirical research into the behavioural and neurological underpinnings of belief, and its evolutionary and prehistoric roots, Ward delves into the domain of what the Berkeley psychologist John Kihlstrom has termed the 'cognitive unconscious' (Kihlstrom 1987; see also Kahneman 2011). A rich array of non-conscious mental activity, including learned responses that have become automatic, subliminal perceptions that impact on our conscious judgements, and implicit but not consciously recalled memories—all these profoundly affect how we perceive and interpret the world (Ward 2014: 11, 68). And as we saw in the case of Freud (whose general influence is clearly discernible here), the implications of the resulting conception of human belief and understanding extend far more widely than the domain of the pathological. Not just in neurotic desires and perceptions, but whenever we believe anything at all, there is, as Ward puts it, a 'mode of liminal processing, related to embodiment and affectivity, which "thinks" more quickly and reacts more instinctively than our conscious rational deliberation' (Ward 2014: 12).

A further dimension of complexity in our beliefs is explored in the neuroscientist Iain McGilchrist's groundbreaking work *The Master and His Emissary*, according to which there are two different modes of relating to the world, broadly correlated with the activities of the left and right hemispheres of the brain respectively, one mode being detached, fragmented, abstract, and analytical, the other being more direct, holistic, intuitive, and empathetic. McGilchrist speaks of:

> … two ways of being in the world, both of which are essential. One is to allow things to be *present* to us in all their embodied particularity, with all their changeability and impermanence and their interconnectedness, as part of a whole which is forever in flux … The other is to step outside the flow of experience and 'experience' our experience in a special way: to *re-present* the world in a form that is … is abstracted, compartmentalised, fragmented, static … From this world we feel detached, but in relation to it we are powerful (2009: 93).

Acknowledging McGilchrist's influence, Graham Ward urges us to question the 'left-brain hegemony' that has increasingly dominated our culture since the Enlightenment, and to reconfigure our understanding of belief. Doing justice to the full range of our embodied human engagement with the world could allow for a 'rebalancing' of left-brain and right-brain modes of awareness. This, he argues, might enable us to overcome the sterile opposition between scientific and religious modes of thinking, and to understand what lies 'at the very core of poetic and religious faith' (Ward 2014: 110).

Before proceeding, it needs to be noted that many critics have questioned the distinction that is invoked here between 'left-brain' and 'right brain' activity, objecting that the available scientific research on the neurophysiology and functioning of the brain does not support a strong dichotomy between the functions performed by the two hemispheres (Nielsen et al. 2013). There may be evidence to suggest that in most subjects one can distinguish between 'logical–conceptual' and more 'intuitive' mental activity, each broadly correlating with neural activity in the relevant halves of the brain; but the critics point out that in normal subjects both halves play some role in both, and in any case there is constant interaction between the two hemispheres. All this, however, is readily conceded by McGilchrist, who fully acknowledges the massive degree of interconnectivity in the wiring of the brain, while nevertheless insisting that the two hemispheres have been shown to function in ways that are to some degree independent, and that this can tell us something important about the different ways in which we experience the world (McGilchrist 2018).[10]

However that may be, the position taken by McGilchrist, Ward, and others[11] about the need to challenge what they term 'left-brain hegemony' does not seem ultimately to hinge on the precise details as to how the brain is configured. For the crucial point at issue is not a neurological one, but what might be called a psycho-ethical or spiritual one: that our ultimate flourishing as human beings depends on our being able to integrate our detached and analytic modes of relating to the world with our more direct and intuitive modes of awareness. This is not to say, however, that the scientific study of the brain has no relevance to the psychological-cum-moral task of striving for an integrated vision of the world. For the wiring of the brain, shaped by the long history of its evolution, is an integral part of our nature as biological creatures, and our human ways of perceiving and understanding the world must inevitably be conditioned and mediated by that history. The point was in fact explicitly anticipated by Jung in a paper written early in his career:

> Just as the human body represents a whole museum of organs, with a long
> evolutionary history behind them, so we should expect the mind to be organized in
> a similar way … We receive along with our body a highly differentiated brain
> which brings with it its entire history, and when it becomes creative it creates out
> of this history—out of the history of mankind … that age-old natural history which
> has been transmitted in living form since the remotest times, namely the history of
> the brain structure
>
> (Jung 1918: 12).[12]

In short, whatever scientific consensus is eventually reached with regard to the precise
workings of the brain and the functioning of its parts, the resulting picture seems likely
only to reinforce the idea that our grasp of reality depends at the physiological level on an
intricate nexus of mechanisms and processing systems evolved over many millennia and
working beneath the threshold of conscious awareness and control. And alongside this
neurological complexity there also has to be taken into account the complex array of
socially and culturally inherited associations and resonances that condition our cognitive
and emotional responses to the world, again working largely below the level of our
explicit conscious awareness. So the more we learn about all this, the more pressure is
put on the idea of the detached autonomous agent, somehow operating above the fray of
evolution and history, and forming beliefs based solely on dispassionate scrutiny of the
evidence like some pure disembodied intelligence.

So what are the implications of the growing interest in the 'archaeology of belief' for our
understanding of religion and its place in the modern scientific age? Ward, as already
noted, maintains that a greater understanding of what lies beneath the surface of
conscious belief-formation will help us to overcome what he sees as the 'sterile'
opposition between scientific and religious thought and to give up the idea that primitive
mythological ways of thinking about the world will progressively be replaced by modern
scientific methods. For placing the belief-forming faculties of our species within the
context of their biological and social development over many millennia reveals the
ineradicable role of the mythic and the symbolic in *all* human cognition, and thus radically
undermines the idea of the inevitable triumph of a science-based, demythologized and
secularized belief system. Just as Jung had argued that mythical and symbolic forms
powerfully and inescapably impinge on our human beliefs and attitudes, so Ward argues
that all human belief systems involve myth-making. And this includes not just archetypal
stories of our origins (such as the Genesis narrative), but a whole range of human activity
—the 'symbolic realms we hominids have been cultivating for 2.2 million years', including
art, poetry, rite, and dance. In ways we cannot fully explain, these interlocking modes of
human culture tap the powers of what (for a want of a better term) we call the
imagination, which operates at many more levels than are accessed by our conscious
reflective awareness. Such works of the imagination 'intimate that our experience … of
being in the world is freighted with a significance that only an appeal to the mythic can
index' (Ward 2014: 186).

Many important issues are raised by this stress on the psychological depth and complexity of human cognition, and its mythical and the imaginative aspects. But for present purposes two key questions present themselves: first, what are the implications of all this for the epistemic status of religious belief; and second, what lessons emerge for the way in which the philosophy of religion should be conducted? To these questions we shall briefly turn in the fifth and final section of our discussion.

Philosophizing About Religion and the Layers of the Human Psyche

The line of argument canvassed in the previous section—emphasizing the creative, imaginative, and mythical elements in all human belief systems—might seem to offer a kind of protective armour for religious ways of thinking against the advances of modern scientific rationalism. But there may be grounds for concern that such a defence of religion is bought at too steep a cost—the cost of eroding the very distinction between *mythos* and *logos*, between imagination and reason. For even if science is necessarily the creation of our evolved human capacities, conditioned by our long human history, it has nevertheless developed tried and trusted methods (empirical investigation, mathematical modelling) for understanding and predicting the workings of nature. And the secularist charge against religious ways of thinking is that they completely fail to pass these tests for reliable belief formation, and thus do not deserve a place in our modern world view.

There is no space here to delve further into the extensive and continuing contemporary debate about the future of religion in the modern world. What needs to be addressed in the present context are the implications for philosophy, and in particular the philosophy of religion, of the issues raised in the previous section about the complex 'archaeology of belief'. In this connection, one does not have to sign up to a questionable assimilation of science and myth in order to wonder if contemporary analytic philosophy of religion has become too dry and austere, too closely modelled on the pared-down unambiguous language of the sciences, to do justice to the complexities of religious belief and the ways in which it might contribute to our understanding of ourselves and the nature of the reality of which we are a part. It is here that the contribution of psychoanalytic thought seems particularly relevant. For if there is any truth in what the psychoanalytic movement has tried to uncover about the hidden layers of the human mind, then it seems plausible to suppose that being more open to 'what lies beneath' might lead to a more nuanced epistemology, less modelled on the austere language and methods of the physical sciences, but arguably better equipped for the task of philosophizing insightfully about religion.[13] It is striking in this connection to find even a committed analytic philosopher of religion such as Eleonore Stump arguing recently that in order to do its job philosophy of religion may require deeper and richer resources than those afforded by the tools of logical analysis and technically expert argument (Stump 2010: 26–7).

Stump does not explicitly mention or invoke the resources of psychoanalytic theory, but significantly she does argue that philosophers of religion need to make use of our manifold responses to the multiple resonances of literary, and scriptural, narrative. This chimes in with earlier calls for a certain kind of narrative or literary turn in philosophy, as advocated for example in the work of Martha Nussbaum. Nussbaum argues that in learning to appreciate a great literary text we have to allow ourselves to be receptive and 'porous', knowing when to yield instead of maintaining constant critical detachment (1990: 281–2). Some philosophers may suppose that any departure from complete analytical detachment would involve a loss of philosophical integrity; and certainly there is need for philosophical caution whenever our imaginative and emotional resources are made use of. But equally, if we insist on maintaining a detached analytical stance at all times, this may be less a sign of intellectual integrity than what Nussbaum calls 'a stratagem of flight' (Nussbaum 1990: 268)—a refusal of the openness and receptivity that is prepared to acknowledge all the dimensions of our humanity.

If this is right, then one lesson to emerge is that we may need a new epistemology for thinking about religious belief and its basis. In contemporary analytic philosophy of religion, both the advocates of religious belief and its critics tend to operate with an *epistemology of control*. We stand back, scrutinize the evidence, retaining our power and autonomy, and pronounce on the existence or otherwise of God. But such methods implicitly presuppose that the divine presence ought to be detectable via intellectual analysis of formal arguments or observational data. Yet the ancient Judaeo-Christian idea of the *Deus absconditus* (the 'hidden God')[14] suggests a deity who is less interested in proving his existence or demonstrating his power than in the moral conversion and freely given love of his creatures, and in guiding the steps of those who 'seek him with all their heart', in Pascal's phrase (1670: no. 427). And when we start to think about the means of such conversion, it becomes clear that it could never operate through detached intellectual argument alone, or through the dispassionate evaluation of 'spectator evidence' (Moser 2008: 47).

Any suggestion that religious claims cannot fully and properly be evaluated from a detached and dispassionate standpoint may at first seem to be special pleading on behalf of religion; but further reflection makes it clear that there are all sorts of other areas of life—appreciation of poetry, of music, entering into any kind of personal relationship— where we need to be (to use Nussbaum's term) 'porous'. Otherwise, while we pride ourselves on being in control and judiciously evaluating the evidence, we may actually be closing ourselves off from allowing the evidence to become manifest to us. In short, there may be many areas of human life where a proper understanding of what is going on requires us to relinquish the epistemology of control and substitute an *epistemology of receptivity*.[15]

Perhaps surprisingly, this plea for the adoption of an epistemology of receptivity when assessing the claims of religion can draw some support from the writings of the founding father of psychoanalysis. Although as we have seen, Freud himself was a stern critic of religion, and although he tended to present himself very much in the garb of the austere

scientific rationalist, he also acknowledged, perhaps most explicitly in his *Introductory Lectures on Psychoanalysis*, that his methods of treatment diverged very significantly from those which are typical of standard scientific procedure. He points to two important differences. First, conventional scientific medicine looks to 'establish the functions and disturbances of the organism on an anatomical basis, to explain them in terms of chemistry and physics, and to regard them from a biological point of view', whereas Freud concedes that his own approach focuses on a much more elusive aspect, namely the 'psychological attitude of mind'. And second, Freud acknowledges the oddity of the fact that the processes involved are not susceptible of public investigation under normal observer conditions, because the psychotherapeutic process takes place in a private consulting room and 'only under the conditions of a special affective relationship to the physician' (1916–17: 22).

Thus, so far from there being an objective scientific template to which all valid discourse and all legitimate human inquiry must conform, Freud in the *Introductory Lectures* appears ready to allow that there are phenomena whose nature is such that quite different modes of understanding are appropriate. Indeed, he goes further and acknowledges that psychoanalysis is learnt first of all 'on oneself, by studying one's own personality' (1916–17: 23). These concessions are most significant, since (whether Freud himself drew such an inference or not) they implicitly cut the ground from underneath those critics of religion who would dismiss the validity of religious experiences on the grounds that they resist external verification by detached or non-involved observers.

The important lesson to emerge here is that despite the prevalence of scientistic modes of thinking in our contemporary culture (and in some parts of Freud's own thinking), we need to take seriously the idea that there may be phenomena that do not manifest themselves 'cold', as it were, but require involvement and commitment on the part of the subject in order to be apprehended. It has been an assumption of modern scientific inquiry that the truth is simply available for discovery, given sufficient ingenuity and the careful application of the appropriate techniques, and that the dispositions and moral character of the inquirer are entirely irrelevant. But while this assumption may be correct enough when inquiring into truths within meteorology, say, or chemical engineering, it seems quite out of place when we are dealing with certain central truths of our human experience—for example truths about how a poem or symphony may be appreciated, or how a loving relationship may be achieved and fostered. In these latter areas, the impartial application of a mechanical technique is precisely the wrong approach: the truth yields itself only to those who are already to some extent in a state of receptivity and trust.[16], [17] The upshot is that there may be phenomena, or parts of reality, whose detection or apprehension is subject to what might be called *accessibility conditions*: the requirements for getting in touch with them include certain requirements as to the subjective attitude and psychological (and perhaps moral) state of the subject (Cottingham 2005: ch. 1, §§ 3 and 4; ch. 7, §4). And as we have seen, Freud himself seems clearly to acknowledge this when he speaks of the insights arising from the psychotherapeutic process making themselves available 'only under the conditions of a special affective relationship to the physician' (1916–17: 23). Yet once it is granted that

there are psychological truths that may come to light only given certain affective and other transformations within the experiencing subject, then it may become possible to see how the same principle might be applied to religious truths, so that certain transformations in the subject may be crucially necessary preconditions for the manifestation of the divine reality that is the object of the religious quest (Cottingham 2009: ch. 5).

What thus emerges, as we bring to a close our discussion of the relation between psychoanalysis and religion, is the remarkable degree of convergence that obtains between these two very different ways of thinking about the human condition. Not only do both outlooks search for deeper layers of significance beneath of the surface world of factual assertion and plain 'common sense', but also, as we have just seen, both hold that this deeper world may disclose itself only to those who are in a suitable state of receptivity—a point that carries important epistemological implications perhaps not yet fully assimilated in our contemporary philosophical culture. To be sure, none of this is sufficient on its own to constitute a vindication of the claims either of psychoanalytic theory or of traditional religion, nor is it intended to be; but at least it may give some indication of the way in which those claims might have to be assessed.[18]

References

Augustine of Hippo (*c.* 398 **CE** [1912]). *Confessions* (*Confessiones*), trans. W. Watts. Cambridge, MA: Harvard University Press.

Cottingham, J. (1998). *Philosophy and the Good Life: Reason and the Passions in Greek, Cartesian and Psychoanalytic Ethics*. Cambridge: Cambridge University Press.

Cottingham, J. (2005). *The Spiritual Dimension*. Cambridge: Cambridge University Press.

Cottingham, J. (2009). *Why Believe?* London: Continuum.

Cottingham, J. (2014). *Philosophy of Religion: Towards a More Humane Approach*. Cambridge: Cambridge University Press.

Cottingham, J. (2015). *How to Believe*. London: Bloomsbury.

Dante Alighieri (*c.* 1310 [1981]). *The Divine Comedy: Paradise* (*La Divina Comedia: Paradiso*), ed. and trans. G. Bickersteth. Oxford: Blackwell.

Dawkins, R. (2006). *The God Delusion*. London: Transworld Publishers.

Foucault, M. (2005). *The Hermeneutics of the Subject*. New York: Palgrave Macmillan.

Freud, S. (1901). *The Psychopathology of Everyday Life*. In J. Strachey (ed. and trans.), *The Penguin Freud Library*, Vol. 5. London: Penguin.

Freud, S. (1916–17 [1952]). *Introductory Lectures on Psychoanalysis* (*Vorlesungen zur Einführung in die Psychoanalyse,*), trans. J. Riviere as *A General Introduction to Psychoanalysis*. New York: Washington Square Press.

Freud, S. (1927). 'The Future of an Illusion'. In J. Strachey (ed. and trans.), *The Penguin Freud Library*, Vol. 12 (pp. 182–241). London: Penguin.

Freud, S. (1929). 'Civilization and its Discontents'. In J. Strachey (ed. and trans.), *The Penguin Freud Library*, Vol. 12 (pp. 251–340). London: Penguin.

Howard-Snyder, D. and Moser, P. (eds) (2002). *Divine Hiddenness*. Cambridge: Cambridge University Press.

Hume, D. (1757). 'The Natural History of Religion'. In *Four Dissertations*. London: Millar.

Jung, C. (1918). 'The Role of the Unconscious' ('Über das Unbewußte'). In C. G. Jung, *Collected Works*, rev. edn (l967–77), Vol. 10. London: Routledge.

Jung, C. (1938). *Psychology and Religion*. In C. G. Jung, *Collected Works*, rev. edn (l967–77), Vol. 11 (pp. 89–95). London: Routledge.

Jung, C. (1939). 'Conscious, Unconscious and Individuation'. In C. G. Jung, *Collected Works*, rev. edn (l967–77), Vol. 9 (1) (pp. 275–289). London: Routledge.

Jung, C. (1951). *Aion*. In C. G. Jung, *Collected Works*, rev. edn (l967–77), Vol. 9 (2). London: Routledge.

Jung, C. (1952). 'Symbols of Transformation'. In C. G. Jung, *Collected Works*, rev. edn (l967–77), Vol. 5. London: Routledge.

Jung, C. (1956). 'Correspondence with H. L. Philp'. In C. G. Jung, *Collected Works*, rev. edn (l967–77), Vol. 18 (pp. 706–707). London: Routledge.

Kahneman, D. (2011). *Thinking Fast and Slow*. London: Penguin.

Kihlstrom, J. (1987). 'The cognitive unconscious'. *Science 237*: 1445–1452.

Knox, J. (2003). *Archetype, Attachment, Analysis*. Hove: Brunner-Routledge.

Leiter, B. (2004). *The Future for Philosophy*. Oxford: Clarendon Press.

Martin, T. (2002). *Oppression and the Human Condition*. Lanham, MD: Rowman & Littlefield.

May, S. (2011). *Love: A History*. New Haven, CT: Yale University Press.

McGilchrist, I. (2009). *The Master and His Emissary*. New Haven, CT: Yale University Press.

Moser, P. (2008). *The Elusive God: Reorienting Religious Epistemology*. Cambridge: Cambridge University Press.

Nagel, T. (1974 [1979]). 'What is it like to be a bat?'. In *Mortal Questions* (ch. 12). Cambridge: Cambridge University Press.

Nagel, T (1994). 'Freud's Permanent Revolution'. *New York Review of Books 41*(9): 34–38.

Nielsen, J. et al. (2013). '**An Evaluation of the Left-Brain vs. Right-Brain Hypothesis**'. *Plos* One *8*(8): e71275.

Nussbaum, N. (1990). *Love's Knowledge*. Oxford: Oxford University Press.

Palmer, M. (1997). *Freud and Jung on Religion*. London: Routledge.

Pascal, B. (1670 [1962]). *Pensées*, ed. L. Lafuma. Paris: Seuil.

Ricoeur, P. (1970). *Freud and Philosophy: an essay on interpretation*, trans. D. Savage. New Haven, CT: Yale University Press (Originally published as *De l'interprétation: Essai sur Sigmund Freud*, 1965).

Sartre, J.-P. (1975). *The Emotions: Outline of a Theory*, trans. B. Frechtmann. Secaucus, NJ: Citadel Press (Originally published as *Esquisse d'une théorie des emotions*, 1939).

Stump, E. (2010). *Wandering in Darkness*. Oxford: Oxford University Press.

Ward, G. (2014). *Unbelievable: Why We Believe and Why We Don't*. London: I. B. Tauris.

Notes:

(1.) Quoted in Martin 2002: 67. It should be noted that Sartre himself was critical of the Freudian concept of the unconscious mind.

(2.) [Eds: See Blass, this volume, § 'The Dialogue Between Freud and the Believer on the Nature of Truth', for an account of Freud's later explanation, in *Moses and Monotheism*, of the force of religious belief in terms of 'historical truth'.]

(3.) [Eds: Cottingham's argument may be fruitfully juxtaposed with Blass, this volume, § 'Contemporary Positive Perspectives' and Boothby, this volume.]

(4.) 'Fecisti nos ad te, et inquietum est cor nostrum donec requiescat in te' (Augustine (*c.* 398), Bk. I, Ch. 1; 'E'n la sua volontade è nostra pace' (Dante Alighieri (*c.* 1310), Canto iii, line 85.

(5.) [Eds. See also Blass, this volume, § 'Contemporary Positive Perspectives'.]

(6.) For more on the psychodynamics of this transformational process, see Cottingham (1998), ch. 4.

(7.) 'Hermeneutics' in the most general sense may be thought of as an approach to philosophy which gives a central place to the (culturally mediated) search for self-understanding. See Ricoeur 1970.

(8.) [Eds. Blass, this volume, § 'The Change in the Attitude to Truth' explores Freud's antipathy to this turn in relation to questions of religion.]

(9.) [Eds. See Gipps, this volume, for a development of this line of thought.]

(10.) See McGilchrist, 'Exchange of Views' at http://iainmcgilchrist.com/exchange-of-views/.

(11.) See for instance Eleonore Stump on 'cognitive hemianopia' (2010: 26–7).

(12.) Compare McGilchrist 2009: 8.

(13.) [Eds. A similar conclusion is reached via a very different route by Blass, this volume.]

(14.) See Isaiah 45:15. For more on the 'hiddenness' of God, see Howard-Snyder and Moser (2002).

(15.) The argument of the last three paragraphs draws on material from Cottingham 2014: ch. 1. See also Cottingham 2015, esp. ch. 3.

(16.) For more on this theme, see Michel Foucault, Seminar at the Collège de France of 6 January 1982, trans. in Foucault (2005: 1–19). See also Cottingham 2005: ch. 5 and ch. 7.

(17.) [Eds. For development of this line of thought in relation to psychoanalysis, see Gipps, this volume.]

(18.) I am grateful to Richard Gipps and Michael Lacewing for their most thoughtful comments on an earlier draft of this chapter.

John Cottingham

John Cottingham is Professor Emeritus of Philosophy at the University of Reading, UK.

Access is brought to you by

Oxford Handbooks Online

Abstract and Keywords

While Freudian psychoanalysis is famous for its negative evaluation of religion, many contemporary analytic schools reject this view, regarding religion positively. Interestingly, both critical and positive psychoanalytic approaches to religion are based on the idea that religious belief is an illusion. This chapter explores these approaches, their development, and the thinking that underlies them especially in relation to the notion of truth and the personal quest for it. It sheds light on the complexity of Freud's thinking on truth and conviction and draws attention to a neglected dimension that is crucial to his understanding of religion. The chapter points to the fact that despite fundamental differences, the Freudian world view has more affinity with that of the religious believer than commonly thought. In turn, the world views of the psychoanalytic schools which have regarded religion positively are actually opposed to those of most believers.

Keywords: Freud, religion, conviction, doubt, truth, psychoanalysis, illusion, spiritual experience, Winnicott, Moses and Monotheism

Rachel B. Blass

(p. 553) Freud's Negative View of Religion

FOR Freud the understanding of any phenomenon is primarily the understanding of the personal motives, or needs and wishes, the phantasies that come into play through it or in relation to it.[1] When it comes to religion Freud's concern is with the individual's motives for belief in God, for faith. And while these would naturally differ from person to person, Freud concluded that across individuals it was possible to discern features common to the state or act of belief in God. These common features determined Freud's evaluation of religion, which was, famously, very negative.

Freud's best-known ideas regarding faith centre on the individual's wish to have a protective father figure with whom he can feel identified. He discussed this idea extensively in his 1927 *The Future of an Illusion* and further elaborated it in his 'The question of a *Weltanschauung*' (1933). In this latter text he summarizes what religion undertakes to do for people as follows:

> It gives them information about the origin and coming into existence of the universe, it assures them of its protection and of ultimate happiness in the ups and downs of life and it directs their thoughts and actions by precepts which it lays down with its whole authority. Thus it fulfils three functions .... [I]t satisfies the human thirst for knowledge; it soothes the fear that men feel of the dangers and vicissitudes of life, when it assures them of a happy ending and offers them comfort in unhappiness ... [and] it issues precepts and lays down prohibitions and restrictions.

> (Freud 1933: 161)

(p. 554) Freud goes on to explain that what unites these three seemingly disparate aspects of religion (instruction, consolation, and ethical demands) is the fact that they are all tied to the child's view of his father. The God-creator whom believers call father, Freud writes, 'really is the father, with all the magnificence in which he once appeared to the small child' (Freud 1933: 163). He created us, he protected us, and he taught us to restrict our desires. Freud explains that when one grows up one still remains helpless in many ways in face of the dangers one encounters in life, but one recognizes that father cannot really be a source of protection from them. Thus, Freud explains, the believer:

> harks back to the mnemic image of the father whom in his childhood he so greatly overvalued. He exalts the image into a deity and makes it into something contemporary and real. The effective strength of this mnemic image and the persistence of his need for protection jointly sustain his belief in God.

> (Freud 1933: 163)

These needs and wishes for the protective father explain not only the idea of the existence of a personal god who created us and loves us, but also our sense of guilt in relation to him. Our feelings of guilt are expressions of our conscience, which we form from critical inner voices of our parents in an effort to be assured of their love. These voices are now perceived as coming from God. Freud concludes:

> The amount of protection and happy satisfaction assigned to an individual depends on his fulfilment of the ethical demands; his love of God and his consciousness of being loved by God are the foundations of the security with which he is armed against the dangers of the external world and of his human environment. Finally, in prayer he has assured himself a direct influence on the divine will and with it a share in the divine omnipotence.

> (Freud 1933: 164)

The notion of sharing in divine omnipotence points to our desire not only for security, but for control; indeed God is great, but God is also in a close relationship with us, He is in us, and like us. This may be a great boost to our ego. In this context, belief in Jesus (the Son) is often regarded as an expression of our desire to be both exonerated of our 'sins' in relation to the father (in sharing the experience on the cross) and a desire to become God ourselves (Freud 1927).

In other words, in the depths of the mind God becomes an idea, a construct, associated with different parts of ourselves and our internal parental figures. He serves us in our desire for protection, for love, for punishment, for restriction of our desires, for exoneration, for perfection and power (while in each person these specific factors may play out somewhat differently). In addition, our need to submit to authority, to follow clear rules, to *not* think on our own, may come into play in the act of faith, both in our relationship to God and in our relationship to those who transmit faith to us. In a sense, Freud here claims that we are naturally inclined to blindly accept traditional claims of belief.

(p. 555) Freud is clear on the fact that having wishes and needs for the world to be a certain way does not mean that the world is *not* that way. And in this context he puts forth his much-cited distinction between 'illusion' and 'delusion'. He writes:

> What is characteristic of illusions is that they are derived from human wishes … In the case of delusions, we emphasize as essential their being in contradiction with reality. Illusions need not necessarily be false—that is to say, unrealizable or in contradiction to reality. For instance, a middle-class girl may have the illusion that a prince will come and marry her. This is possible; and a few such cases have occurred. That the Messiah will come and found a golden age is much less likely. Whether one classifies this belief as an illusion or as something analogous to a delusion will depend on one's personal attitude. Examples of illusions which have proved true are not easy to find, but the illusion of the alchemists that all metals can be turned into gold might be one of them. The wish to have a great deal of gold, as much gold as possible, has, it is true, been a good deal damped by our present-day knowledge of the determinants of wealth, but chemistry no longer regards the transmutation of metals into gold as impossible. Thus we call a belief an illusion when a wish-fulfilment is a prominent factor in its motivation, and in doing so we disregard its relations to reality, just as the illusion itself sets no store by verification.[2]

> (Freud 1927: 31)

While on the basis of this distinction Freud may have concluded that beliefs about God are illusions, he goes on to assert that they are, in fact, delusions; they have no source *but* the psychological one. He describes the rationale behind this assertion as follows: 'scientific work is the only road which can lead us to a knowledge of reality outside ourselves' (1927: 31), and religious ideas are not only closed to scientific investigation

but also are 'so improbable, so incompatible with everything we have laboriously discovered about the reality of the world' (1927: 31).

The fact that religious beliefs are held by many does not alter the situation. In his *Civilization and its Discontents* (1930) Freud speaks of the situation in which large groups 'attempt to procure a certainty of happiness and protection against suffering through a delusional remoulding of reality' and adds that 'The religions of mankind must be classed among the mass delusions of this kind' (1930: 81).

The negative attitude of Freud, of Freudian psychoanalysis, and in fact of a good part of psychology at large as well, to religious faith has to do with the combination of the kind of psychological motives that maintain it and its delusional nature. At bottom, it is regarded as a distortion of reality, of truth, due to self-serving, infantile needs. Thus faith is regarded not only as mistake, but as a kind of moral failure, childishness at best, and, accordingly, inclinations towards it should be overcome. Freud (p. 556) affirms that, when 'a psychologist who does not deceive himself' assesses the process of human development:

> the idea forces itself upon him that religion is comparable to a childhood neurosis, and he is optimistic enough to suppose that mankind will surmount this neurotic phase, just as so many children grow out of their similar neurosis. (1927: 53)

Contemporary Positive Perspectives

Ever since the time these critical views were initially put forth there have been voices coming from within psychoanalysis (in addition to the numerous ones coming from without it) that have taken issue with them. Carl Jung is often presented as Freud's earliest and most stringent opponent in this regard.[3] In contrast to Freud's rejection of religion he went as far as to state that religion is curative. He wrote:

> Among all my patients in the second half of life—that is to say, over thirty-five—there has not been one whose problem in the last resort was not that of finding a religious outlook on life. It is safe to say that every one of them fell ill because he had lost what the living religions of every age have given to their followers, and none of them has been really healed who did not regain his religious outlook.

> (Jung 1933/2001: 234)

What Jung meant by adopting a religious outlook has been the subject of much discussion. And while it is commonly understood as an expression of Jung's belief in God and his appreciation of such belief by others, some have argued that Jung's use of religious language is misleading; that his reference to God is, actually, to certain aspects of the human psyche and that his call is not for an encounter with the divine, but with the self. As Martin Buber states in his critique of Jung's portrayal of modern man: 'Man does not deny a transcendent God; he simply dispenses with Him…. In his place he knows the

soul, or rather the self' (Buber 1952/1988: 86). This explains how Jung, in his book *Answer to Job*, could speak of man helping God attain wholeness, in part through God coming to a better integration of his evil sides.

In 1913 the ties between Freud and Jung were severed and Jung went on to develop his own school of psychological thought and practice, Analytical Psychology, which differed with the Freudian school in many ways. Focusing on development within the latter, the first critical voice of Freud's views on religion is that of Oskar Pfister, the Swiss psychoanalyst and Lutheran minister. Pfister's objections were made public with Freud's blessing in his paper 'The Illusion of a Future', published in German in *Imago* in 1928 (and republished in English in the *International Journal of Psychoanalysis* in 1993). It is (p. 557) possible to follow a continuous, albeit rather thin, stream of critical writings that ensued in the following fifty years. Zilboorg, Erikson, Fromm, and Loewald are the most prominent of the writers who explicitly rejected what was considered to be Freud's central view of religion as illusionary, delusionary, infantile, and neurotic. The more dominant trend, however, was the acceptance of this Freudian view (Kakar 1991: 56). From the 1980s, a change in this trend began to take place and a wave of religion-friendly writings emerged. Among some of the more provocative book titles directly addressing the issue of psychoanalysis and religion that have appeared since then are *The Birth of the Living God* (Rizzuto 1979), *The Psychoanalytic Mystic* (Eigen 1988), *The Analyst and the Mystic* (Kakar 1991), *Soul on the Couch* (Spezzano and Gargiulo 1997), *Emotion and Spirit* (Symington [1994] 1998), *Ecstasy* (Eigen 2001), *Terror and Transformation* (Jones 2002), accompanied by some more low-key titles such as *Psychoanalysis and Religious Experience* (Meissner 1984) and *Psychoanalysis and Religion* (Smith and Handelman 1990). In these books, and in the numerous articles that have appeared during these years, religion is understood in the light of psychoanalytic theory (and especially through some of its recent developments) as a normal, healthy, positive phenomenon. The understanding expressed in this literature reflects a broader sentiment within the psychoanalytic field in general of tolerance and acceptance of religious belief and believers (as may be seen in David Black's (2006) collected essays *Psychoanalysis and Religion in the 21st Century: Competitors or Collaborators?*).

Under the influence of Freud's thinking on religion, analysts could describe a patient's loss of religious belief as a measure of success (Fenichel 1938: 316) and continued religious faith and practice as a sign of failure (1951: 189, cited in Leavy 1990: 50). Today, while analysts are as atheistic as they were in Freud's day, they would not express such views (even if they privately held them). A dramatic change has taken place.

One may point to four main lines of thinking that characterize the new views on religion and to some extent explain their emergence:

    1) Freud failed to appreciate the value of illusion.

According to the newer literature on religion, when the meaning of illusion is fully grasped the question of the validity of religious belief that seemed to trouble Freud becomes irrelevant. Not only is there no possible move from illusion to delusion, in the

sense that the fact that religious belief is founded on a wish does not allow one to conclude that the belief is false, but also the fact that a belief is an illusion implies the irrelevance of the question of falsity. Here the thinking of the British psychoanalyst D. W. Winnicott on the early development of the child, published primarily in the 1950s and 1960s, plays a crucial role. For example, Winnicott describes the importance of the infant believing that he created the breast that feeds him, and the importance of allowing him to maintain this illusion. This is clearly false in a simple factual sense, but in another sense one may say that the breast arrives because of the infant's scream. He also describes how the infant becomes invested in objects (his thumb, his blanket) as though they were the breast, and in a certain sense they take on its meaning. This situation is, according to (p. 558) Winnicott, neither real nor imaginary nor delusional, but rather involves some transitional area of reality, one of illusion. It's an area in which the question: ' "Did you conceive of this or was it presented to you from without?" ... is not to be formulated' (1953: 14).

Winnicott goes on to suggest that this understanding of the nature of illusion and its role in infant development has broader cultural implications as well, including for religion. In his paper 'Transitional Objects and Transitional Phenomena' (1953) he describes how the mother allows:

> the infant the illusion that what the infant creates really exists. This intermediate area of experience, unchallenged in respect of its belonging to inner or external (shared) reality, constitutes the greater part of the infant's experience, and throughout life is retained in the intense experiencing that belongs to the arts and to religion. (1971: 16)

This perspective has been very much taken up by newer analytic approaches to religion and allows them to ground their positive view of religion as something that exists in a kind of transitional space. Indeed, as Freud says, religion is an illusion, but illusions are necessary and valuable and cannot be dismissed by questioning whether they tell us something true about objective reality.

It should be noted here that the irrelevance of the question of objective reality refers only to the *content* of the illusion. The *capacity* to form illusion and the relational conditions that enable this are considered to be objective realities of great relevance and an important focus of analytic concern.[4] This leads to the next characteristic of the new analytic models of religion.

2) Religion satisfies basic self and relational needs.

The contemporary positive perspective on religion maintains that there is a connection between religion in some senses of the term and psychic well-being. They introduce criteria of well-being, which were not of direct concern to Freud and his followers, to support this. Religion, they argue, may allow for the expression, and even reinforcement, of various desirable human capacities for relationship, such as trust, intimacy, care, and community (Wallwork and Wallwork 1990: 161), and provide an important means for

articulating basic aspects of self-experience. As James Jones, the very prolific American writer on psychoanalysis and religion, states:

> religious moments … allow us to enter again and again into that timeless and transforming psychological space from which renewal and creativity emerge … [there,] we gain access again to the formative (and reformative) experience at the heart of selfhood. (1991: 134)

(p. 559) In presenting such capacities and experiences as relevant factors in determining the value of religion these authors rely on a post-Freudian psychoanalytic approach which is thought to better appreciate the value of regressive forms of relatedness and experience. For example, undifferentiated relatedness to a maternal object is considered necessary for early developmental processes and in order to overcome painful feelings of loss later in life. Thus religious experience which supposedly gives expression to undifferentiated forms of relatedness to a maternal object are thought to be helpful to the individual's well-being (Kakar 1991).

3) Religion is defined in terms of broad spiritual experience.

In the contemporary positive analytic perspective on religion, the ideal religion becomes more of a personal, self-determined mysticism, devoid of history, ritual, authority, obligation, and mediation, a kind of Westernized Buddhism. This religious ideal often does away with the notion of the existence of a transcendent object to the point that the self becomes the object of religious experience. For example, Symington, in his 1994 book on psychoanalysis and religion, defines religion and spirituality so that, in their mature forms, they are compatible 'with the denial of God's existence' (1994: 88) and equated with willing the good, disciplining one's intentions, and upholding a morality based on freedom and responsibility for the other. Coming from a very different angle and referring to the writings of Lacan, Bion, and Winnicott, Michael Eigen speaks of holiness, mysticism, sacredness, and prayer, apparently referring by these terms to a general and vague kind of openness to experiencing (Eigen 1988, 2001). Responding to the question of what he means by 'God' he states:

> God could be anything … to blank oneself out and be totally open to whatever currents pulse this way or that … whether you're into body, or emotion … Taoist or Buddhist, whatever, it feels good. (1988: 193; the last two ellipses appear in the original)

At the same time, a critical stance is often maintained in regard to more traditional forms of religion, which are regarded as dogmatic and fundamentalist. Hinshelwood explains in his foreword to an edited book entitled *Beyond Belief*, which presents contemporary views on the nature of the relationship between psychoanalysis and religion, that he feels a kind of aversion towards religion in its institutionalized form. He adds that he thinks he shares the viewpoint of all the contributors to the volume (twelve in number), who 'more or less all repudiate "organised religion". They reject dogma and (mostly) rituals. They

revert simply to "natural religion", the deeply personal religious experience of a mystical kind' (Hinshelwood 1999: xvi–xvii).

   4) Psychoanalysis is regarded as a spiritual experience.

The contemporary positive analytic perspective on religion tends to regard not only religion in terms of a more general spiritual experience, but psychoanalysis too. Symington, for example, considers the aim of psychoanalysis to be 'the transformation of bad actions into good' and consequently maintains that it is a spiritual endeavour (1994: 181). (p. 560) He concludes that both theology and psychoanalysis have yet to recognize that 'the greatest spiritual encounters' (1994: 130) occur between patient and analyst. Moreover, also being, according to Symington, a practice geared towards the transformation of narcissism into concern for others, psychoanalysis is the pinnacle of mature religiosity and a much-needed substitute for the failed primitive religions of revelation (1994: 75).

Along similar lines Eigen speaks of the 'holiness' of therapy (1988: 42), and the 'mystical' and 'sacred' (2001: 37) nature of psychoanalysis and the self-experiencing that these allow for. He tells us that psychoanalysis is, for him, at times 'a form of prayer' (1988: 11).[5]

The Change in the Attitude to Truth

These four features that characterize the positive view of religion in psychoanalysis point to the fact that those who have rejected Freud's stance on religion do so primarily on the basis of a change of attitude in relation to the role and value of truth. Freud rejected religion because he considered it to involve a distortion of truth. He maintained that it is founded on belief in claims that are false and this in and of itself is not only wrong, but also can never be considered conducive to the person's development and well-being. What psychoanalysis can offer is only cure through truth and aimed at seeing reality as it is (Freud 1917). The positive analytic view of religion argues, in contrast, that through changing the value of illusion, by expanding the scope of its relevance, the question of truth can be bracketed. Moreover, neither psychoanalysis nor religion should be measured in terms of its claims to truth. In place of these we should focus on the personal, interpersonal, and spiritual benefits that each allows for.

In other words, the shift towards a positive view of religion, for the most part, comes with a transformation of both the nature of psychoanalysis and the nature of religion. They begin to resemble each other. Notably, the question most central to the traditional Judaeo-Christian believer, the question of God's existence, is not only bracketed, but is no longer considered relevant or meaningful. Religion is regarded not as an expression of knowledge or truth pertaining to the nature of reality, a transcendent reality of God, but more as a kind of self- or relational experiencing within a realm of illusion. And it is self-

and relational experiencing of this kind which is considered to be what psychoanalysis offers and aims for.[6]

Freud clearly would have rejected such developments in psychoanalysis. Already in his time the view was gaining force that religious beliefs are merely a kind of fiction accepted as true for their great practical significance, both culturally and in terms of the needs of mankind that cannot be fulfilled by 'cold science' (1927: 35). However, such a (p. 561) view, Freud states with clear disdain, could be proposed only by a *philosopher*. No serious believer would accept this. Nor could he. In a letter to Oskar Pfister in 1930 Freud writes, 'The question is not what belief is more pleasing or more advantageous to life, but of what may approximate more closely to the puzzling reality that lies outside us' (Meng and E. L. Freud, 1963: 132–3). Freud's concern and efforts were directed towards coming to know this reality.

As I've noted, Freud's argument against religion did not rest solely on its being based on false claims. He was also concerned with its infantile origins—a point much emphasized by Freud's critics. But it may be seen that these infantile origins only complement, support, and contribute to Freud's position that religious beliefs are not actually true. In effect, by pointing to religion's infantile roots Freud explains how it is that people believe such beliefs, given that it cannot be because of the actual existence of the realities that these beliefs purport to be about. However, it is the truth and not the infantilism that is the crucial point here for Freud. Clearly other activities that are infantile in origin are not intolerable to Freud. Adult sexual activity also has infantile roots but, as long as it does not make false claims regarding reality, Freud did not feel that it should be overcome. Similarly, art, although expressing phantasy life and turning away from reality, could be viewed as valuable because it made no claims to universal truth.[7] 'Art', Freud writes, 'does not seek to be anything but an illusion … it makes no attempt at invading the realm of reality' (1933: 160). It was only the invasion of reality, of truth, that was intolerable to Freud, and which he thought to be taking place in the case of religious belief.

Prophetically, perhaps, Freud knew that his position on the truth of religion would not be welcomed by modern defenders of the faith. He remarked that, as a consequence of his stance, he would now 'have to listen to the most disagreeable reproaches for my shallowness, narrow-mindedness and lack of idealism or of understanding for the highest interests of mankind' (1927: 36). He 'consoled' himself with the thought that at least in our day and age such reproach would not bring with it 'a sure curtailment of one's earthly existence and an effective speeding-up of the opportunity for gaining a personal experience of the after-life' (1927: 36).

While Freud has indeed been scorned by believers, it would seem that Freud's view that believers are making claims to truth would be better received by traditional believers than the more religion-friendly analytic views that have emerged in recent years. While all believers would be happy to hear that being a believer does not necessarily entail involvement in infantile, regressive, neurotic processes and practices, and on the contrary could reflect and contribute to one's psychic well-being and adaptive living

within society, for many the essence of religion lies in acknowledging the actual reality of a supernatural or divine being (Leavy 1990: 47, 50; Vergote 1990: 76). These believers maintain that there is no meaning or value to religion if its assertions are not true, not real. 'Real' here does not mean real in the sense of being expressive of real inner experience, (p. 562) a psychic reality that lies behind the religious stories and practices, but rather real in the sense that the assertions made about the nature of God, his existence and transcendence, his actions and his promise, his message and his demands, refer to something actual. The Judaeo-Christian tradition in this context speaks of the existence of a unique being who actually intervenes in the course of human history. As the theologian Hans Kung summarizes:

> The man who believes … is primarily interested … in the reality itself … He wants to know whether and to what extent his faith is based on illusion or on historical reality. Any faith based on illusion is not really faith but superstition. (1984: 418)

Cardinal John Henry Newman expressed similar sentiments. He writes that 'religion, as a mere sentiment, is to me a dream and a mockery. As well can there be filial love without the fact of a father, as devotion without the fact of a Supreme Being' (*Apologia pro vita sua*, chapter II [1864] 1995: 39). What is affirmed here is not belief in the literal truth of all biblical stories (e.g. that Jonah spent three days in the belly of a fish), but in the truth of a God whose existence and nature finds expression through these stories.

In sum, Freud rejected the truth of religious belief and thus the value of religion and, in so doing, took a stance diametrically opposed to that of traditional believers. However, it may be argued that in that very opposition he was closer to the concerns of the traditional believer than those who adopt the much more positive attitude characteristic of the new psychoanalytic thinking on religion. This is because Freud's opposition, his rejection of religion, is ultimately founded on a love of truth. As Freud rejected religion in the name of truth, he not only maintained that religion consisted of assertions that made truth claims, but he also shared in the passionate desire for truth with which the traditional believer, in his faith in an ultimate reality, would readily identify.

Oskar Pfister, who wrote against Freud's views on religion, thought that this love of truth pointed not only to an affinity with the believer, but was, in fact, a religious stance. Shortly after the publication of Freud's most critical text on religion, *The Future of an Illusion*, Pfister wrote in a reply to Freud that 'Anyone who has struggled so powerfully for the truth and fought so courageously for the redemption of love as you have is, whether he admits it or not, a true servant of God according to scripture' (Pfister 1928: 149–50; in Meissner's translation, 1984: 75).

The Dialogue between Freud and the Believer on the Nature of Truth

Recognizing that Freud and traditional believers share the common ground of a concern with and love of truth, further exploration of that ground is possible.

> 1) Freud's recognition of his inability to make assertions regarding religious truth and yet his insistence on doing so.

(p. 563) Freud openly acknowledged that psychoanalysis can determine what is real or not only in relation to psychic reality—its area of expertise. It can tell us about wishes and other kinds of personal motivations that come into play *in relation* to other areas or aspects of reality, but it cannot make pronouncements regarding the actual nature or existence of that reality. As noted earlier, understanding that we wish for reality to be a certain way doesn't tell us about the actual nature of reality. To use Freud's terminology, the fact that we have illusions doesn't tell us that they are delusions (that they contradict reality) (Freud 1927: 31). For example, while psychoanalysis may tell us of the motivations involved in the acceptance of the law of gravity or the theory of relativity, it does not inform us of the truth or falsity of such laws and theories, nor does it add new physical laws of its own. Freud's critics emphasize his failure to apply this understanding to the reality of God's existence as well.

This criticism is legitimate. While Freud held that the determination that a belief is delusional could not be made on the basis of psychoanalytic thinking or findings, in regard to religious belief Freud moves quite quickly from illusion to delusion, offering but a few extra-analytic considerations as to why he thinks such belief contradicts reality. Freud's personal and cultural reasons for his digression from his own position when it comes to religious belief have been the subject of much discussion (e.g. the role and sources of his own atheism). It may, however, be seen that there is an intrinsic analytic one as well. This has to do with the supernatural status of religious truth, which renders its claims about the basic nature of reality beyond objective demonstration. Indeed psychoanalysis' concern is only with coming to know psychic reality, but it may be argued that the possibility of doing so relies on having a conception of the basic nature of external reality. For example, how we would understand the psychic reality of a patient who denies the law of gravity is strongly influenced by what we know or assume regarding the existence and acceptance of this law and what we expect of the patient in this regard. Similarly, how we understand the psychic reality of a patient who affirms the existence of God will be shaped by basic conceptions regarding the reality of God's existence (see Leavy 1990: 59). Since there is no consensus or objectively demonstrated position on this basic conception of reality, when patient and analyst differ in regard to this conception, that is, when they differ in their views of religious truth, the search for psychic truth may be constricted. For Freud such constriction of the analytic project was unacceptable and this may be seen to underlie his firm stance that religious truth was not only an illusion, but that it directly contradicted what is known about reality (1933: 160).[8]

2) Historical truth: Some similarities between the form and content of psychoanalytic and religious truth.

Moses and Monotheism, which was ultimately published in full as a book in 1939, is often viewed as yet another of Freud's anti-religious texts (for example, Gay 1987: 44; Meissner 1984: 106). In that book, Freud explains the origins of Judaism and monotheism through a very complex analysis of religious and historical writings. Freud's explanation (p. 564) centres on a story regarding the relationship between Moses, an Egyptian prince and follower of the monotheistic beliefs of Akhenaten, and the problematic Jewish people whom Moses chose in order to carry on his religion. This relationship was ultimately a fatal and painful one, as the Jews for various reasons and needs were to rise up and kill Moses, only to feel afterwards remorse and guilt. These feelings are reminiscent of the experiences of the sons in Freud's account of the early development of mankind (which he describes in his Totem and Taboo 1913), after they rise up and kill the revered and beloved primal father. Freud's story of this relationship and the historical events surrounding it is presented as a product of careful scrutiny of relevant texts, and as a valid explanation of the source of many of the particular characteristics of the Jewish people, their beliefs, and the attitudes that they have aroused and continue to arouse in others.

In the course of this exposition, Freud presents his notion of 'historical truth'. He does so in the context of his concern, and I have argued that this is the major concern of the book (Blass 2003, 2004), to understand the source of the believer's conviction, the power of certain ideas to compel us to believe in them even though they are not only unverified by reality, but are incompatible with it. Freud wonders, for example, why such ideas are not 'listened to, judged, and perhaps dismissed, like any other piece of information from outside' (1939: 101). In dealing with this question in *Moses and Monotheism*, Freud more seriously than ever before is willing to examine the pious religious suggestion that these ideas, though strange and unverified, are accepted because they are simply true (1939: 128–9). That is, the believer is convinced because he meets in these ideas what he somehow knows to be true; much in the same way that we believe there is a blue chair in front of us for the simple reason that it so happens that this is the case. For example (and this is Freud's main example), the believer accepts the idea of there being a *single* God, although it is a strange and unverifiable proposition, because as a matter of fact there is a single God. The believer accepts the idea because, upon hearing it, the believer's mind recognizes its truth. Freud acknowledges that such an explanation does have explanatory power but, of course, it would require us to also accept that the human mind has a 'special inclination for recognising the truth' (1939: 129). Freud concludes that, knowing that our intellect is very easily led astray by our wishes—that is, we tend to take as true that which we wish for—the claim that we have such a special inclination for truth is obviously mistaken and consequently the pious explanation must be rejected. Nevertheless, very interestingly at this point, Freud does not suggest (as he has at others) that what the believer recognizes as true is, therefore, merely illusion and in turn delusion. Instead, here the notion of historical truth is introduced. He writes, 'We too believe that the pious solution contains the truth—but the *historical* truth and not the

material truth' (1939: 129). This historical truth, as Freud goes on to describe and define it, does not refer to the historical facts; it is not the objective reality of the past (as it is often wrongly taken to mean). But it is not merely an internal, psychical reality either. Rather, it is a special kind of impression in the mind of a past reality. Its origin is in the real world; real events left the impression. But the impression is not identical to the real world, to the actual events. The source of the relevant distortion in this context is not our wishful tendencies, but rather the primal nature of the impression.

(p. 565) Something from the outside was registered in our minds at a time when it was impossible for us to create an accurate image of it, or to understand it in a way that is accurate in terms of our present understanding of reality. It is this historical truth, buried in the depths of our minds, according to Freud, that explains the believer's acceptance of religious ideas. These ideas are false propositions regarding objective reality, but they are in a sense true to the internal impression of what was once external reality. As such, they are what Freud refers to as 'justified memor[ies]' that demand to be believed (1939: 130). In other words, the believer is compelled to accept religious ideas because upon hearing them his mind recognizes their truth, but what takes place here is not a simple correspondence between mind and reality as such (as when my mind recognizes the existence of a chair that is there), but rather between mind and the vestiges of a past primal reality. To the extent that religious ideas are shaped by the limitations of our mind to have accurately perceived and registered this reality, and, to the extent that they are further influenced by our wishes, these ideas are false. However, to the extent that they convey the great truths of 'the earliest experiences of the whole of humanity' that return and find revived expression in our minds and in our lives, they are also true—true to a past external reality (1939: 130).

Here there is a radical shift in Freud's position regarding the truth of religious ideas and their relationship to his own. Indeed, Freud had always maintained that there exist some real events at the foundations of religious belief, whether it be a primeval patricide or a loving relationship that did, in fact, exist in the past with one's own father. But now there is a difference. I mention here just two points. First, Freud's focus is now on the truth and the justification of these ideas, not on their distortive nature. There are good and real reasons to believe. As he notes in the postscript to *An Autobiographical Study* (written while he was preparing his *Moses and Monotheism*): 'In *The Future of an Illusion* I expressed an essentially negative valuation of religion. Later, I found a formula which did better justice to it … granting that its power lies in the truth which it contains' (1935: 72). Second, while religious ideas are distortive of the objective material reality, Freud now, in effect, contends that *all* ideas that attempt to grasp the common primal realities of mankind will inevitably be so. It is in the very nature of that early reality that we encounter it only through the traces that it leaves in our mind. We can then retrospectively form propositions regarding the material reality that engendered these traces, but these propositions are always only conjecture. Our minds have never really fully met that early reality, nor, being prehistorical and external in origin, could we ever meet it now. And while these propositions are only conjecture they are still the best we

have in the way of truth, since it is through these conjectures, these convictions, that the traces of an earlier reality can now find some expression (Blass 2003: 676).

It is important to recognize that in *Moses and Monotheism* Freud not only expresses ideas on the origins of religion and its value, but also offers a psychoanalytic perspective on the person and on knowledge, which in important ways corresponds to those of religious approaches. To explain: (a) Freud's position in *Moses and Monotheism* that belief in God is not merely a projection of various aspects of the father of childhood, (p. 566) but reflects true and valuable aspects of early human reality, highlights an aspect of Freud's basic perspective on the person that is often overlooked and is downplayed by Freud himself.

According to this perspective the person's foundational psychic reality has an inherent ethical texture to it (see Blass 2006, 2012). That is, Freud's description of this reality is not a neutral portrayal of psychological facts, of what happens to exist in our mind; rather his description of the facts contains an inbuilt prescription on how the person should be. It tells us not only that people develop Oedipal conflicts and tend to feel guilty, but rather that in experiencing certain forms of guilt we are actually fulfilling our human nature, giving expression to 'historical truth'—and it is right that we do so. And here we may see that the nature of historical truth that Freud thinks the person should give expression to, should come to know and live, corresponds in important ways to what is foundational to the traditional believer's view of the person. The foundational reality that Freud thinks the person must contend with centres on a battle between love and hate, between conflicted desires, which results in destruction and specifically in the killing of the beloved father, which arouses remorse and which accounts for a basic sense of guilt that guides us throughout life. Of course, despite this kind of correspondence, Freud rejects the believer's view that this nature is the result of the intervention of an actual transcendent God, and here lies a fundamental and very important difference.

A similar approach to the foundational reality of the person may be seen in basic ideas later put forth by the British psychoanalyst Melanie Klein and her followers. According to the Kleinian view psychic reality is composed of early relationships between objects that are 'good' and/or 'bad'. But this designation of good and bad does not merely refer to qualities of feelings. For example, the mother's breast is not only *experienced* as good; it is good not only as a source of pleasure and nutrition. Rather the breast is regarded as an actual and direct source of goodness. It reflects goodness that exists in the world as well as in the person's internalized psychic reality. Accordingly, the ability to recognize this goodness rather than deny it (e.g. out of envy) is what it means to see truth and to properly live one's psychic truth: (b) Freud's recognition of the value of historical truth puts in question the criteria for accepting a claim as true or delusional that, as we have seen, he argues for elsewhere (Freud 1927, 1933).

On the basis of this recognition he accepts claims as true even where there is no way of objectively demonstrating their truth. Moreover, he suggests that a sense of conviction in the truth of the claims, like the believer's belief, can serve as an indication—at least in

part—of their truth. The conviction, while possibly delusional, a distortion of truth based on wishful thinking, may also be evidence that we have encountered foundational truths regarding our very nature, truths so primary, so integral to human nature, that they cannot be corroborated by material evidence. Their truth, rather, is lived and is made apparent through the very sense of conviction itself.

But with this new approach to conviction based on the notion of historical truth, there comes the danger that there would be no room for doubting convictions, as the sense of conviction itself might be taken as a sign of truth. Freud's solution is to maintain (p. 567) a balance between accepting conviction and questioning one's wishful motives for accepting it, a balance between belief and doubt. Interestingly, one can see this tension even in the very writing of *Moses*. Throughout that text Freud struggled with his doubts as he continued to seek material truth and proof, but ultimately he allowed himself to go with his convictions (Blass 2003). Noting the lack of sufficient evidence for his conclusions, he confides to Andreas-Salomé that 'It suffices me that I myself can believe in the solution of the problem' (letter of 6 January 1935, cited in Pfeiffer 1966: 205).

It may be seen that this more complex notion of truth and how we come to know it, with its focus on lived convictions and an ongoing process of doubting them, better captures the clinical stance that Freud recommends than Freud's more familiar objective approach (Blass 2016). The psychoanalytic process is not about proving to the patient facts about what's going on in his mind, but of encountering, and coming to know and question and better integrate, one's inner reality. It's a fluid process in which the sense of conviction regarding what one comes to understand plays an important part.[9]

Dialogue and Tension

As I have described, contemporary positive psychoanalytic views of religion bridge psychoanalysis and religion by affirming the existence of a common ground of illusion, by putting aside distinctions between truth and fiction, religious ideas and those of psychoanalysis, experience of the sacred and experience of the self.

Freud, in contrast, opposed religion; he did not accept the traditional belief in God's actual existence or in anything supernatural. Only human reality lies at the foundation of his explanation of our experiences and tendencies. But his opposition was grounded in a concern to encounter truth. The person, he maintained, should strive to see reality as it is rather than distort it in order to make ourselves more loved, protected, forgiven, and ideal than we are. Freud's concern for truth is shared by traditional believers. Moreover, in his later writing additional areas of commonality become apparent as his notion of truth becomes increasingly complex. According to his more complex notion foundational truths of our mind are not regarded as something to be simply objectively demonstrated or proven. Rather, conviction and doubt play an important part in their discernment. And the view of the person and his basic predicament that finds expression in these truths—

his ethical struggle with love, destruction, and guilt towards a paternal figure—significantly resembles aspects of certain religious views of the person.

References

Black, D. M. (ed.) (2006). *Psychoanalysis and Religion in the 21st Century: Competitors or Collaborators?* London: Routledge.

Blass R. B. (2003). 'The puzzle of Freud's puzzle analogy: Reviving a struggle with doubt and conviction in Freud's *Moses and Monotheism*'. *The International Journal of Psychoanalysis 84*: 669–682.

Blass, R. B. (2004). 'Beyond illusion: Psychoanalysis and the question of religious truth'. *The International Journal of Psychoanalysis 85*: 615–634.

Blass, R. B. (2005). 'Attachment and separateness and the psychoanalytic understanding of the act of faith'. In J. Auerbach, K. Levy, and C. E. Schaffer (eds), *Relatedness, Self-Definition and Mental Representation: Essays in Honor of Sidney J. Blatt*. New York: Routledge.

Blass, R. B. (2006). 'The role of authority in grounding and concealing truth, "Truth and Tradition in Psychoanalysis" '. *The American Imago 63*: 331–353.

Blass, R. B. (2012). 'Sin and transcendence vs. psychopathology and emotional well-being: On the Catholic Church's problem of bridging religious and therapeutic views of the person'. *Spiritus: A Journal of Christian Spirituality 12*: 21–42.

Blass, R. B. (2016). 'The quest for truth as the foundation of psychoanalytic practice: A traditional, Freudian-Kleinian perspective'. *Psychoanalytic Quarterly 85*: 305–337.

Blass, R. B. (2017). 'Du fantasme de séduction à la foi: la lutte de Freud avec le doute et la conviction au sujet de la vérité de ses idées comme prémisse d'une approche psychanalytique de la connaissance'. In A. Abella and G. Dejussel (eds), *Conviction, Suggestion, Séduction*. Paris: PUF.

Blass, R. B. (2018). 'Commentary on "Lynn Kuttnauer's Venturing onto Thin Ice: Tiptoeing in the Wake of Empathic Failures" '. In R. Tuch and L. Kuttnauer (eds), *Conundrums and Predicaments in Psychotherapy and Psychoanalysis: The Clinical Moments Project* (pp. 64–68). New York: Routledge.

Bloom, A. (1987). *The Closing of the American Mind*. New York: Simon & Schuster.

Bori, P. C. (1979). 'Una pagina inedita di Freud: La premessa al romanzo storico su Mosè' [An unknown page of Freud: The premise to the historical novel of Moses]. *Rivista di Storia Contemporanea 7*: 1–16.

Buber, M. (1952/1988). *Eclipse of God: Studies in the Relation between Religion and Philosophy*. Atlantic Highlands, NJ: Humanities Press International.

Deutsch, H. (1951). *Psychoanalysis of the Neuroses*. London: Hogarth Press.

Eigen, M. (1988). *The Psychoanalytic Mystic*. New York: Free Association Books.

Eigen, M. (2001). *Ecstasy*. Middletown, CT: Wesleyan University Press.

Fenichel, O. (1938). 'Problems of psychoanalytic technique'. *Psychoanalytic Quarterly* 7: 303–324.

Freud, S. (1910). 'Leonardo da Vinci and a memory of his childhood'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 11 (pp. 57–138). London: Hogarth Press.

Freud, S. (1913). *Totem and Taboo*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 13 (pp. vii–162). London: Hogarth Press.

Freud, S. (1917). 'A difficulty in the path of psychoanalysis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 17 (pp. 137–144). London: Hogarth Press.

Freud, S. (1927). *The Future of an Illusion*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 21 (pp. 1–56). London: Hogarth Press.

Freud, S. (1930). *Civilization and its Discontents*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 21 (pp. 57–146). London: Hogarth Press.

Freud, S. (1933). 'The question of a *Weltanschauung*'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 22 (pp. 170–197). London: Hogarth Press.

Freud, S. (1935). 'Postscript to *An autobiographical study*'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 20 (pp. 71–74). London: Hogarth Press.

Freud, S. (1939). *Moses and Monotheism: Three Essays*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 23 (pp. 1–138). London: Hogarth Press.

Gay, P. (1987). *A Godless Jew: Freud, Atheism and the Making of Psychoanalysis*. New Haven, CT: Yale University Press.

Green, A. (2001). *Time in Psychoanalysis: Some Contradictory Aspects*. London: Free Association Books.

Hinshelwood, R. D. (1999). 'Foreword'. In S. M. Stein (ed.), *Beyond Belief: Psychotherapy and Religion* (pp. xv–xix). London: Karnac Books.

James, W. (1902). *The Varieties of Religious Experience: A Study of Human Nature*. New York: Longmans, Green and Co.

Jones, J. W. (1991). *Contemporary Psychoanalysis and Religion: Transference and Transcendence*. New Haven, CT: Yale University Press.

Jones, J. W. (2002). *Terror and Transformation: The Ambiguity of Religion in Psychoanalytic Perspectives*. East Sussex: Brunner-Routledge.

Jung, C. G. (1933/2001). *Modern Man in Search of a soul*. Abingdon: Routledge Classics.

Kakar, S. (1991). *The Analyst and the Mystic: Psychoanalytic Reflections on Religion and Mysticism*. Chicago, IL: Chicago University Press.

Kermode, F. (1985). 'Freud and interpretation'. *International Review of Psychoanalysis 12*: 3–12.

Kristeva, J. (1987). *In the Beginning was Love: Psychoanalysis and Faith*. New York: Columbia University Press.

Kung, H. (1984). *On Being a Christian*. New York: Image Books.

Laplanche, J. (1995). 'Seduction, persecution, revelation'. *The International Journal of Psychoanalysis 76*: 663–682.

Leavy, S. (1990). 'Reality in religion and psychoanalysis'. In J. H. Smith and S. A. Handelman (eds), *Psychoanalysis and Religion* (pp. 43–59). Baltimore, MD: Johns Hopkins University Press.

Meissner, W. W. (1984). *Psychoanalysis and Religious Experience*. New Haven, CT: Yale University Press.

Meng, H. and Freud, E. L. (eds) (1963). *Psychoanalysis and Faith: The letters of Sigmund Freud and Oskar Pfister*. New York: Basic Books.

Newman, J. H. ([1864] 1995). *Apologia pro vita sua [A defence of his life]*. New York: Penguin Books.

Ostow, M. (1988). 'Psychoanalysis and religious experience'. *Journal of the American Psychoanalytic Association 36*: 206–209.

Pfeiffer, E. (ed.) (1966). *Sigmund Freud and Lou Andreas-Salomé: Letters*. New York: Harcourt, Brace, Jovanovich.

Pfister, O. (1928). 'Die Illusion einer Zukunft' ['The illusion of a future']. *Imago 14*: 149–184. (p. 570)

Richardson, W. (1992). 'Love and the beginning: Psychoanalysis and religion'. Contemporary Psychoanalysis *28*: 423–441.

Rizzuto, A. (1979). *The Birth of the Living God: A Psychoanalytic Study*. Chicago, IL: Chicago University Press.

Smith, J. H. and Handelman, S. A. (1990). *Psychoanalysis and Religion*. Baltimore, MD: Johns Hopkins University Press.

Spezzano, C. and Gargiulo, G. J. (eds) (1997). *Souls on the Couch: Spirituality, Religion and Morality in Contemporary Psychoanalysis*. Hillsdale, NJ: Analytic Press.

Symington, N. ([1994] 1998). *Emotion and Spirit: Questioning the Claims of Psychoanalysis and Religion*. London: Karnac Books.

Vergote, A. (1990). 'Confrontation with neutrality in theory and praxis'. In J. H. Smith and S. A. Handelman (eds), *Psychoanalysis and Religion* (pp. 74–94). Baltimore, MD: Johns Hopkins University Press.

Wallwork, E. and Wallwork, S. (1990). 'Psychoanalysis and religion: Current status of a historical antagonism'. In J. H. Smith and S. A. Handelman (eds), *Psychoanalysis and Religion* (pp. 160–173). Baltimore, MD: Johns Hopkins University Press.

Winnicott, D. W. (1971). *Playing and Reality*. London: Penguin Books.

Yerushalmi, Y. H. (1991). *Freud's Moses: Judaism Terminable and Interminable*. New Haven, CT: Yale University Press.

Notes:

(1) [Eds: See also Cottingham, this volume, § 'Freud's Critique of Religion'.]

(2) In this chapter I will describe psychoanalytic positions on concepts of illusion, delusion, truth, reality, and objectivity as presented by their authors. Within this single text it would not be possible to subject them to philosophical critique and this task is left to the readers.

(3) [Eds: See also Cottingham, this volume, § 'Carl Jung and the Importance of Symbolic Thought'.]

(4) It should also be noted that not all writers who have promoted the notion of illusion and transitional space for the understanding of religion have dismissed the objectivity of the content of the illusion. Notable exceptions are William Meissner and Ana-Maria Rizzuto. However, others have made use of their writings to support the general trend.

These ideas regarding the necessity of illusion are, at times, accompanied by the idea of the necessity and value of the irrational (for example, Ostow 1988: 209).

(5) This may be seen to be part of a broader cultural trend, which has allowed for a new respect for the sacred by changing the meaning of the term (see Bloom 1987: 216).

(^{6}) [Eds: See Cottingham, this volume, for the development of a view that seeks to avoid both poles of this disagreement.]

(^{7}) This understanding stands in contrast to some of the newer analytic interpretations, which maintain that Freud held a generally antagonistic attitude towards illusion, except in the realm of art, and that, were this attitude to change for the better, so would Freud's attitude towards religion (Kakar 1991: 59; Meissner 1984: 162).

(^{8}) For another way of thinking about Freud's approach to dealing with claims regarding religion's truth within the analytic situation see Blass 2018.

(^{9}) [Eds: See also Cottingham, this volume, § 'Philosophizing about Religion and the Layers of the Human Psyche' on similarities between religion and psychoanalysis.]

Rachel B. Blass

Blass Heythrop College, University of London, UK

Oxford Handbooks Online

The No-Thing of God: Psychoanalysis of Religion After Lacan 🔓

Richard Boothby

The Oxford Handbook of Philosophy and Psychoanalysis
Edited by Richard G.T. Gipps and Michael Lacewing

Abstract and Keywords

This chapter traces some main lines of Jacques Lacan's interpretation of religion and divinity, which differs significantly from Freud's critique. Orienting ourselves with respect to what Lacan calls das Ding, the enigmatic desire of the Other, it is possible to sketch a Lacanian analysis of religion parallel to that offered by Kant in Religion Within the Limits of Reason Alone. The difference is that where Kant looked to find in religious representations, especially those of Christianity, the underlying dynamics of pure rationality and of a morality founded upon it, Lacan discerns the very structure of subjectivity and its relation to the unknown Other. New perspectives are thereby opened up on a whole series of problems, including the unconscious dynamics of enjoyment, practices of sacrifice, the structural differences between various religions, and Christian doctrines of incarnation, love, and mystical unknowing.

Keywords: Lacan, Real, Imaginary, Symbolic, the Other, das Ding, paternal metaphor, sacrifice, sexuation, mysticism

Introduction

On the topic of religion, it is clear that the French psychoanalyst, Jacques Lacan, doesn't merely repeat the critique mounted by Freud. Not so clear is the meaning of what Lacan says on his own.[1] At different points in his teaching, Lacan associates the ancient gods with the unthinkable Real (2017: 44), he identifies the 'I am that I am' of the burning bush with the pure subject of speech (1968: Session of 11 December), he asserts that 'the true formula for atheism is that God is unconscious' (1981: 59), and he intimates that his *Écrits* should be compared with the writings of the great mystics (1998: 76). That last provocation is sounded in Lacan's twentieth seminar, *Encore: On Feminine Sexuality, The Limits of Love and Knowledge*. It is there, too, that Lacan associates an excessive enjoyment beyond knowing, what he calls the '*jouissance* of the Other', with the essential incompleteness of woman, her characteristic of being *pas tout*, or 'non-all'. He then brings such feminine *jouissance* suggestively but enigmatically into conjunction with God, asking at one point: 'Why not interpret one face of the Other, the God face, as based on feminine *jouissance*?' (1998: 82).

How to make sense of these obscure passages? How are we to grasp the intersection they imply between *jouissance*, language, lack, love, and God? And how are these terms knotted in relation to knowing and what is beyond knowing? While such questions cannot be adequately answered in a brief paper, we might hope to provide a basic orientation to them.

Freud's Critique

Freud's critique of religious belief is familiar and need only be tersely summarized to orient our discussion.[2] In *The Future of an Illusion* (1927), Freud's attack is directed less towards the *object* of belief, the dubiousness of an invisible, supernatural power who oversees human life from a cosmic distance, than towards the *subject* who believes. The problem is the motive of the believer. The 'illusion' at stake is thus a matter of the believer's investment in wish-fulfilling fantasies.

To this bottom-line assessment, Freud added two more properly psychoanalytic angles of interpretation. The first, developed in *Totem and Taboo* (1911) and revisited in Freud's last major work, *Moses and Monotheism* (1939), predicates the solidarity of the human group upon the fantasy of the father's murder. The bond of brotherly love is cemented by the blood of the patriarch. It is a theory that casts a light on the origins of religious ritual in practices of blood sacrifice and suggests that what Catholics call the mystical body of the Church is founded upon a collective guilt over the murder of the Saviour.

Freud based a second strategy on the resemblance of religious ritual to the compulsions of the obsessive neurosis. Not only are they both marked by a conspicuous conscientiousness, they also display the compromise structure of the symptom, which functions to repress some impulse at the same time that it symbolically discharges it. The case of the Ratman illustrates the dynamic perfectly (Freud: 1909: 153–320). His obsessional thinking becomes unmanageable upon hearing a fellow soldier describe an oriental torture in which a cage of starving rats is affixed to the buttocks of the victim so that the rats chew their way into the anus. The Ratman becomes preoccupied with the recurring dread that such horrifying treatment might be applied to the two most important figures in his life: his father and his fiancée. In the frightful shudder that accompanies each compulsive reimagining, Freud recognizes both a struggle to defend against a sexually charged desire and a means of satisfying that very desire. Key religious phenomena might be interpreted in parallel fashion. Contemplation of the crucifixion, for example, invites the worshipper to rise above all hatred in a final victory of love while also offering an occasion to gratify an unconscious sadism.

Lacanian Heresy

In a simple but elegant pun, the title of Lacan's twenty-second seminar, 'R.S.I.,' ––which pronounced in French sounds like 'heresy'––suggests both a measured departure from Freudian orthodoxy and, more subtly but palpably, a promise of a new angle of interpretation on the problem of religion. The three letters stand for the three fundamental registers––Real, Symbolic, and Imaginary–– with which Lacan rewrites the essentials of psychoanalytic theory. The case of the Ratman offers a useful frame in which to introduce them.

The fantasy of bodily violation in the rat torture recalls the dynamics of the 'imaginary', which for Lacan inform the structure and function of the ego. During the period of late infancy that Lacan calls 'the mirror stage', the unifying gestalt of the other's body is thought to coordinate the impulse chaos of the child. This reliance on an external image alienates the human subject from itself. It inevitably invites aggression towards the other because it introduces an internal split. On this point, Lacan rests a central postulate of his entire outlook: the opposition between the *moi* and the *je*, between the *ego* as the seat of one's conscious sense of self, and the *subject* of unconscious desire.

That subject emerges in pulsations of the 'symbolic', Lacan's general term for the resources of language, informed by the theories of Saussure and Jakobson, on which he draws in order to retrieve and radicalize Freud's many analyses of word play in dreams, slips, bungled actions, and symptomatic complexes. In the Ratman, for example, the obsessive idea clearly crystallizes around the morpheme 'Rat', which functions like a kind of switch-point in a constellation of emotionally loaded issues: the starving, gnawing rodents (*Ratten*), monetary debts (*Raten*), gambling (*Spielratte*), and marriage (*Heiraten*). The key thing that Lacan adds to Freud's own elaborate analyses (and what he adds to the linguistic theories of Saussure as well) is a more general theory of the signifier centred on its unconscious dimension. The signifier always passes into the speech stream an indeterminate excess, what Lacan calls the '*objet petit a*'. In the margin of this excess there is a 'slippage of the signified beneath the signifier' (Lacan, 2006: 419) The consequence, put in simplest terms, is that we always say more than we intend. It is in the space of the unsaid and unintended that unconscious effects unfold.

By locating the roots of unconscious desire in the play of the signifier, Lacan reconceives the Oedipus complex as a function of the acquisition of language.[3] What Freud had tied to paternal threats of castration becomes a matter of tensions internal to the child's submission to the symbolic law. Lacan's third category of the 'real' arises at the limit of the law, or better, at the point where the symbolic function remains incomplete, interrupted, or inconsistent. In the gaps and failures of the symbolic, the subject may be confronted with a traumatic excess, the subjective impact of the real that Lacan calls *jouissance*. The real in this sense, far from being reducible to everyday 'reality', indicates the subject's encounter with something fundamentally unthinkable. If it is an encounter with something incomprehensible in the outside world, it is also an engagement with the

incomprehensible in the subject itself. Here, too, the Ratman provides us with a suggestive example. In talking about the dread rat torture, Freud tells us, the patient's face 'took on a very strange, composite expression. I could only interpret it as one of *horror at pleasure of his own of which he himself was unaware*' (Freud 1909: 166–7).

Where, then, are we to locate the religious in the architecture of Lacanian theory? For many commentators, the answer is to identify God with an aspect of the unthinkable real. Divinity is to be situated in a domain beyond human experience, a zone that forever outstrips all capacities to image or name it.[4] Such an approach offers the satisfaction of locating the divine in a dimension of radical transcendence, but it also runs the risks attendant upon the conception first put forward by Kant, who famously claimed to limit human reason in order to make room for God. For such a view, the divine is vouchsafed by being located in a noumenal Beyond, but only at the price of our being barred from saying anything whatever about its nature.

A second angle of interpretation takes its point of departure from Lacan's conception of the symbolic 'big Other'. Part of what is distinctive about Lacan's conception of language is that he characterizes the network of the symbolic code as susceptible of being construed as something akin to, or in some way inhabited by, another subject. This so-called 'big Other' is the nameless and faceless regulator who oversees the written and unwritten rules that direct our lives. At stake is not just the properly linguistic rules governing grammar, syntax, and semantics. The big Other also monitors our polite behaviour, inspects our adherence to fashion, defines for us what is funny (sometimes actually laughing for us in the 'laugh track' of TV sitcoms), insists on our silent decorum at one moment and calls for our deep-throated patriotic fervour at another. The power of this symbolic big Other resides in part precisely in the fact that it exists everywhere and nowhere, that it is both 'all and none'.[5]

In this supposition of something subject-like in the symbolic code we rightly discern a shadow of Freud's primal father. Yet we also get a hint that, unlike Freud, who posed faith as a means of coping with human helplessness by means of pleasing illusions, for Lacan the disposition to belief in God is more emphatically unconscious and more definitely structural. Far from being an illusory add-on to a rationality that could and should do without it, an implicit 'faith' in the underlying coherence of sense-making is a kind of structural requirement of our relation to language. The Lacanian view thus suggests that something like faith, the baseline credulity that undergirds reliance on all symbolic conventions, is a virtual necessity for human beings. It is this function that underlies the promotion of what Lacan calls 'master signifiers', the legitimacy of which are warranted by the big Other in the role of a 'subject supposed to know'. And what is God but the ultimate form of such supposition? If we mortals fail to grasp the meaning of our lives, the same is not true of the divine agency whose plan for Creation is complete in every detail.[6] 'The subject supposed to know', says Lacan, 'is God, full stop, nothing else' (1968: Session of 30 April).

God in the real or in the symbolic. Must we make a choice? A prime purpose of the remainder of this chapter is to show how the two perspectives are in fact inseparable. The key for making their connection lies in Lacan's concept of *das Ding*, the Thing.

In the Shadow of the Thing

While mentioned only infrequently after the central role it plays in the Lacan's seventh Seminar on '*The Ethics of Psychoanalysis*', the concept of the Thing is clearly fundamental.[7] As Lacan says of it, '*Das Ding* is a primordial function which is located at the level of the initial establishment of the gravitation of the unconscious *Vorstellungen*' (1992: 62). How, then, to interpret it? Much of the commentary on Lacan's notion has followed a Kantian clue. On this account, the Lacanian Thing is akin to the Kantian *Ding-an-sich*, the unthinkable kernel of external things, not merely as they appear to us, but as they are in themselves. It is along this line of thinking that Lacan's Thing is associated with the original lost object that sets in motion the ceaseless roving of human desire. *Das Ding* also names the dimension of what is fundamentally excessive (and thus is connected in Lacan's thought to the *object petit a*). As such, it animates the experience of the sublime, which Lacan famously characterizes as emerging when 'the object is elevated to the dignity of the Thing' (1992: 112).

This quasi-Kantian posing of the Thing in the locus of the *Ding-as-sich* is by no means without some value, but it needs to be resisted at least long enough to recognize that the crucial dimension of the *das Ding* concerns not objects but other people. The original unthinkable object is the fellow human being. This conclusion is obvious when we return to the text of Freud from which Lacan takes his clue. In a brief flight of theorizing in his unpublished 'Project for a Scientific Psychology', Freud notes how the child divides the figure of the *Nebenmensch*, the 'neighbour' or fellow human being, between what it can recognize on the basis of similarities to its own body––precisely the sort of mirror recognition that Lacan associates with the imaginary—and a locus of something that is 'new and non-comparable', a zone of something unknown. This unrepresentable excess Freud calls *das Ding*. It is this division of the human other, a division that reserves in the heart of the familiar a locus for something excessive and as yet unknown, that will serve as the template for all of the child's future attempts to interrogate the nature of objects. 'For this reason', says Freud, 'it is in relation to the fellow human-being that a human-being learns to cognize' (1886–99: 331) The key point is not merely that there remains a inaccessible, noumenal core of all objects but that the original schema that locates such an unknown is taken from the example of what is unknown in the human Other.[8]

What is new and crucially important in Lacan's treatment of the Thing is the way in which the enigmatic locus of something uncognized in the Other becomes the root source of anxiety. 'Not only is [anxiety] not without object', he says, 'but it very likely designates the most, as it were, profound object, the ultimate object, the Thing' (Lacan 2014: 331). The challenge of the neighbour-Thing consists not simply in the discovery of an inaccessible kernel at the heart of the Other but in the way it raises the unsettling question of what object I am for that unknown desire. The question presses with particular force in the drama of toilet training when the demand of the Other for the regulation of the infant's bowels re-energizes anxiety about the unanswered question of

what the Other wants. The Lacanian thesis thus goes beyond merely locating the source of anxiety in the fellow human being. It asserts, contrary to our fondest myth about childhood, that the hidden source of deepest and most uncanny anxiety is the mother herself.[9]

We might remark in passing how this Lacanian understanding of the dynamics of the mother/child relation at least partly resembles that offered by Simone de Beauvoir. At a key point of her argument in *The Second Sex*, de Beauvoir reasserts the value of the Freudian Oedipus complex for understanding the deep roots of ambivalence towards the feminine, though with the crucial proviso that we invert Freud's conception (de Beauvoir 1989: 195–6). The core of the Oedipus complex is not, as Freud thought, that the child must be separated from the mother by a threat of castration, but rather that the child is motivated to navigate its own separation, seeking to achieve an autonomy that can be won only by a certain rejection of the maternal embrace. Lacan echoes this key point. He could well be paraphrasing de Beauvoir when he insists that 'it's not true that the child is weaned. He weans himself. He detaches himself from the breast' (Lacan 2014: 313).

Does Lacan then follow de Beauvoir here? By no means completely. Their point of convergence only makes it more essential to clarify Lacan's distance from de Beauvoir, who departs from Freud merely in claiming that the child's desire for the mother is abrogated not by the threat of paternal castration but rather by the child's own rejection. For Lacan, by contrast, the problem isn't the desire of the child at all, but rather that of the mother, in as much as her desire is encountered as a threatening unknown. It is in the light of this perspective that we can make sense of Lacan's comparison of the mother to the daunting spectre of a giant praying mantis (Lacan 2014: 22). In the same stroke, we can interpret his characterization of the *objet petit a* as *un objet cessible*, a cedable or yieldable object (see Lacan 2014: 312–15, 324–7). In the various incarnations of the *objet a*, most of them parts of the body, the subject's 'pound of flesh', is exchanged or 'sacrificed' in order to create a margin of safe separation from the Other.[10] Lacan calls to mind a similar function when he likens the mother to a crocodile and the phallic signifier to a stick with which to keep its jaws from snapping shut.

The notion of the *objet cessible* recapitulates in Lacan's own theoretical frame something akin to Heidegger's notion of an escape from anxiety made possible by means of focusing on an object (Heidegger 1962: 234). What Lacan calls *objet a* ultimately derives from a pure void, a pure function of lack and absence. Yet this lack is successively given definite shape in a series of objects, in the first instance parts of the body (breast, faeces, penis etc.).[11] In this way, the yawning lack discovered in the unfathomable opacity of the Other our topic, the result is a new angle of view on the meaning of sacrifice. In its ultimate function, 'sacrifice is not at all intended to be an offering, nor a gift, both of which are propagated in a quite different dimension, but the capture of the Other in the web of desire' (2014: 277).

This discussion of the materialization of the *objet a* and its deployment in an economy of exchange provides a ready segue to the child's entry into language, the most basic and highly elaborated economy of exchange. In the originary drama with the maternal Other, the inarticulate cry of the infant becomes in itself a ceded object, indeed the very first such object, yielded into the space between subject and Other. As Lacan says, 'this manifestation of anxiety coincides with the very emergence in the world of he who is going to be the subject. This manifestation is his cry … this first effect of cession … the nursling can't do anything about the cry that slips out of him. He has yielded something and nothing will ever conjoin him to it again' (2014: 326).

If the first inchoate eruptions of the voice are in this way inflected with anxiety, and inevitably so, in as much as they enter and begin to shape the emerging interval between the subject and the Other, they are also destined to become the means by which the question of the Other, the enigma of *das Ding*, will be ceaselessly reposed. 'Here we're touching', says Lacan, 'on the very thing that makes the relation to the Other possible, that is, on that whence emerges the fact that there is such a thing as a signifier. This site whence emerges the fact that there is such a thing as the signifier is, in one sense, the site that cannot be signified. It is what I call the site of the lack-of-signifier' (2014: 134). What is at stake is the most elemental implication of Lacan's theory of the signifier, perhaps precisely what he had in mind by claiming that he had defined the signifier as no one else had dared. In every entry into language, in every iteration of signifying material, there resounds some echo of the unanswered and unanswerable question of the Other.

This perspective on *das Ding*, in which the primordial question of the Other resonates–– waveringly, uncertainly––every signifier, illuminates the point that Lacan never tires of emphasizing: that it is always possible to ask about anything spoken, 'yes, I hear what you're telling me, but what are you really saying (what do you really want)?'. What most distinguishes Lacan's view of language and its function is that meaning can never be fully stabilized, that a question not only can but always implicitly *is* posed by every entry into language. Yet if every play of the signifier implicitly resumes an interrogation of the unknown in the Other, what is ultimately at stake is the subject's own coming-to-be, the subject's own question. This even more elusive stake of the game is rooted in the real of the subject's mute *jouissance*. It is a locus that can be approached only mythically, conceivable only as a primordial, indeterminate X. The key point for Lacan, however, is that the question of the Other always comes first, that the question of the subject's desire can be posed only in the locus of the Other.[12] 'What anxiety targets in the real'. Lacan says, 'includes the *x* of a primordial subject moving towards his advent as subject, […] since the subject has to realize himself on the path to the Other. […] [this subject] is the subject of *jouissance* […] it can in no way be isolated as a subject, unless mythically' (2014: 173).

Hell is Other People

The upshot of the foregoing discussion, as shockingly violent to common sense as it is to mainstream linguistics, is to assert that the most archaic function of language, far from connecting the subject to the fellow human being, is to achieve an indispensable degree of separation from the Other, to establish a margin of detachment and independence that puts the neighbour-Thing at a distance.[13] The primordial accomplishment of the signifier is to provide a stand-in for the unknown dimension of the Other, putting that threatening unknown at a remove and thereby forestalling the outbreak of anxiety.

With this process in view, we can say with fresh literalness that 'the word is the murder of the Thing'. The phrase refers to the philosophy of Hegel.[14] Yet if the Hegelian formula now becomes meaningful in a new way, another Hegelian concept becomes even more relevant, insofar as the function of the signifier becomes an exemplary instance, indeed *the* exemplary instance of Hegel's cardinal concept of *Aufhebung*, which forms the inner pivot of his famous dialectic.[15] In accord with the two-fold meaning of the German word, something that undergoes *Aufhebung* is simultaneously *cancelled and preserved*. In exactly such a dialectical movement, the signifier both cancels *das Ding*, distancing the subject from it, offering an arbitrary sound in the place of the unknown Thing, yet also preserves precisely what is cancelled, marking it for further coginizing sometime in the future. The signifier establishes a locus suspended between the subject and the Other, at once putting what remains unknown and potentially threatening about the Other at a safe distance, yet thereby repositioning it and maintaining it as a question.

This perspective integrates the most basic points of Lacan's theoretical framework. To identify the elementary function of the signifier with an *Aufhebung* of the enigmatic Thing in the Other is nothing other than to repose the basic terms of Lacan's 'paternal metaphor', his way of reconceptualizing the Oedipus complex, in which a signifier, what Lacan calls the 'Name of the Father', is substituted for the unknown Desire of the Mother. In the same stroke, it explains how and why the inauguration of the signifier is already and in itself the most primitive inscription of the law against incest. There can be no subject without opening a certain distance from the maternal Thing.

To summarize the view we're adopting here, the very being of the subject is predicated upon a suspension of the unknown reality of the Other, the continued management of which becomes the charge of the symbolic law. This perspective is audible in a question put forward by Slavoj Žižek, apropos Lacan's distance from Levinas: 'What if the ultimate function of the Law', Žižek asks, 'is not to enable us not to forget the neighbor, to retain our proximity to the neighbor, but, on the contrary, to keep the neighbor at a proper distance, to serve as a kind of protective wall against the monstrosity of the neighbor? (2005: 163). From this perspective, the lack in the Other, the enigma of the Other's

jouissance, becomes the undetectable 'dark matter', by definition invisible to ordinary consciousness, that underlies the entire edifice of the social world.

We can now return to the question of religion and God with new resources. I'll limit myself to sketching four brief points, by necessity more suggestive and programmatic than comprehensive.

Economies of Jouissance

The analysis we have pursued allows us to trace the basic outline of a Lacanian interpretation of the religious in a manner reminiscent of Kant's (1793) effort in *Religion Within the Limits of Reason Alone*, highlighting those features of religious practice and belief that dovetail with features of human life discernible by wholly secular analysis. The difference is that where Kant's approach was oriented towards the human capacity for reason, for Lacan the crucial problem is the very structure of subjectivity, a structure, as we have seen, fundamentally oriented towards the desire of the Other. Where Kant remains in the orbit of *cogito*, Lacan is attuned to *desidero*. The crucial dimension is not thought but desire, not rationality but *jouissance*.

In Seminar 16––the seminar entitled '*D'un Autre à l'autre*' ('From an Other to the other') —Lacan offers a particularly pithy note about how the big and little Others are to be distinguished, a note that reminds us that the problem of the unknown in the Other is always a question of an unknown enjoyment. The little Other, the fellow human being or neighbour, is the locus of an incipient *jouissance*, the big Other is a second locus that has been somehow cleared, or evacuated, of enjoyment.

> This neighbor, is it what I have called the [big] Other, what I make use of to make function the presence of the signifying articulation in the unconscious? Certainly not. The neighbor is the intolerable imminence of enjoyment. The [big] Other is only its cleared out *terreplein*. […] It is precisely that, it is a terrain cleared of enjoyment
>
> (1968: Session of 12 March).

This passage points to an elemental shift in which an oppressive and intrusive presence of enjoyment in the little Other is traded for some kind of 'clearing', or leveling of enjoyment in the big Other. The implication is that the move towards a reliance on a big Other of the symbolic code answers to the challenge of the neighbour-Thing by means of 'taming' the desire of the Other. We're invited to see the so-called 'clearing' of enjoyment as a structural moment, a kind of necessary base camp or staging, on the way to the establishment of a regulated economy of enjoyment. What sort of regulation is this and how does it work?

The key to moving forward from this point and making new sense of the question of religion and God is to recognize that Lacan's reference to a 'clearing' or evacuating of enjoyment in the big Other doesn't mean that the locus of the big Other remains forever empty of *jouissance*. On the contrary, Lacan refers repeatedly to the big Other as a special site of projected enjoyment. The big Other comes to be another locus, another scene, of enjoyment but one that bears a special relation to the structure of the Law. From this vantage point, religious belief presents us with the spectacle of a congregation of human beings, gathered in the common cause of controlling too raw an exposure to one another's enjoyment in favour of devotion to another, imagined subjectivity in the form of a deity who controls access to enjoyment in accordance with sacred dictates.[16]

Such an economy of *jouissance* is shown with particular clarity, and with particular relevance to questions of the religious, in Lacan's conception of masochism and sadism. His primary point is to correct the vulgar conception of the masochist as someone who simply enjoys pain and the sadist as one who enjoys inflicting it. What is crucial, Lacan insists, is the way in which both masochism and sadism are triangulated by reference to the locus of the big Other. The masochist submits to pain in service of the larger purpose of inciting the anxiety of the big Other. The sadist, for his part, seeks to make himself the instrument of the big Other's enjoyment.

This rereading clarifies a great deal, not least the theatricality of masochism and sadism, the tendency to elaborately *stage* torments. The ultimate spectator of perversion is the gaze of the big Other. Among other things, this means that the results of the perverse game for the little Others—the perverse couple who enact it—are typically not only different but virtually the opposite of what they are for the big Other. In the case of the masochist, far from inciting the anxiety of the torturer, the masochist's submission offers to the sadistic partner a thrill of power, a complete escape from anxiety (the beating scenes from Lars Von Triers' film *Nymphomaniac* illustrate the dynamic perfectly). The anxiety is produced not in the human Other, but wholly on the level of the big Other. It is the big Other whose rules are violated, who is scandalized by pain, injury, and death. Likewise, in sadism, the crucial enjoyment is not that of the torturer (who, as we see in Sade's endless scenarios of torment, is virtually mechanical, unfeeling) but the big Other, who enjoys in his place. In the tenth Seminar, Lacan explicitly makes the link to the religious question: masochism and sadism are devoted respectively to the anxiety and the *jouissance* of God (Lacan 2014: 163). The conclusion to be drawn would have been familiar to St Augustine. Insofar as perversion is always played out under the gaze of the big Other, the perverse is coincident with the dimension of the religious.

From Mortals to Divinities ... and Back Again

For Lacan, the initial separation of the child from the mother is achieved not under pressure of a paternal threat of castration, but by means of the emerging agency of language. The relevance of this point for a new psychoanalytic interpretation of religion is compactly signalled by the opening line of John's Gospel: 'In the beginning was the Word'.

It is now possible to expand upon that mapping of the primal origin of the subject by emphasizing the role of the signifier as an *objet cessible*, a 'ceded object' injected into the space between the subject and the Other. This mediating object can be related to what is arguably the most elemental gesture of all religious life: the act of sacrifice. The signifier is itself the primal instance of something given up by the subject in order to found an economy of exchange with the Other. The sacrificial offerings of religion are essentially signifiers whose function is to establish the field of mediation between subject and Other. [17]

As we have seen, that process of signification inevitably tends towards unfolding of a third locus, that of the symbolic big Other, that grounds the consistency of the code, thereby stabilizing the subject's relation with the little Other of the fellow human being. It is with reference to this triangular structure of the subject, the little Other, and the big Other that Lacan is able to explain the source and strength of the disposition to belief in gods. Where Freud starkly opposed the blindness of faith to all rationality, Lacan roots the motive sources of belief in the elemental architecture of subjectivity. At the most fundamental level, the appeal to deities is grounded in the necessity for stabilizing the referential valence of linguistic signifiers. [18]

Reference to the elementary triangle of subject, little Other, and big Other also promises a means of explaining the logic of a certain evolution of religions, a sort of phylogenesis of religious life, moving across the trajectory from polytheism to monotheism and displaying a marked tendency towards collapsing the hypotenuse of the triangle that separates the little and big Others.

At one extreme, we have something like Aztec sacrifice, centred on an absolutely inhuman Other (the sun) that requires unending human sacrifice. The divine being is a sort of fearsome cosmic parasite that feeds ceaselessly on spurting human blood.

In ancient polytheistic paganism, the gods still require copious sacrifice though a new, much more advantageous arrangement is arrived at in which, as Hesiod tells us, humans get to eat the slaughtered victims and the gods content themselves with the smell of roasting meat wafting towards the heavens. In pagan sacrifice, specific objects are exchanged between mortals and divinities, stabilizing a link between them but also maintaining an unbridgeable distance.

In Judaism, the deity becomes a single individual and the dynamic of sacrifice alters radically. To seal the covenant with Abraham, Yahweh wants only the foreskin, thus reducing the stakes of sacrifice to an almost purely symbolic exchange. The focus on such a relatively trivial object, a conspicuously expendable part of the body, albeit one attached to the privileged organ of vitality, indicates that the value of sacrifice is no longer to be measured by the number or value of victims offered but entirely by the willingness of the sacrificer to give it up. The upshot is to signal an interest less in the

object of sacrifice than in the inward intention, the allegiance and obeisance, of the one who sacrifices.

To this basic view, the commentary of Slavoj Žižek adds an interesting sidelight. Žižek argues that the Hebrew prohibition of idolatry should be interpreted as a response to the fact that the deity has become so close to being human that there arises a kind of anxiety of coincidence, a fear, in effect, that the god is too close for comfort. In the terms we've been forging, this proximity threatens to collapse of the distance between the little and big Others, in turn undermining the economy of enjoyment that is predicated upon it. The problem with graven images and idols is not that they fail to do justice to the transcendent majesty of God but rather, as Žižek puts it that they 'would render [the Jewish God] too faithfully, as the ultimate Neighbor-Thing' (2001: 130–1).

Yet it is precisely such an identification of the god with the neighbour-Thing that is proposed by Christianity. The Christian 'fulfilment of the law' is achieved through a radical reversal of the order of sacrifice, a reversal that is, as Paul says, folly for the Greeks and scandal for the Jews: God sacrifices to save man. Even more radically, God becomes man. The Christian incarnation thus completely collapses the distance separating the big Other and the little Other. The new God can now say to his disciples: 'As you do to the least among you, you do unto me'. 'I am there wherever two or more of you are gathered'.

Sexuation and the Sacred

If the function of the religious centrally involves establishing an economy of *jouissance* that is triangulated between the little and big Others, then we can expect the essential structures and internal tensions of religious life to reflect Lacan's theory of sexuation, by which he sought to distinguish masculine and feminine forms of *jouissance*.[19] To see how they do, we need a little more background.

Among the most important and influential features of Lacan's late teaching are his so-called 'formulae of sexuation', his gloss on Freud's famous question about 'what does a woman want?'. His scheme derives from a deliberate torsion––many will say simply 'distortion'–– of classical logic, resulting in two different modes of relating to the universal. The first, characteristic of the masculine position, defines the universal by reference to the exception. As Freud's formulation had it, all men are subject to castration only because one man, the primal father, wasn't. For the masculine subject, the law is founded upon the point of its transgression. The masculine organization thus tends to be oriented towards prohibition and the fantasy of its exceptional violation. The feminine position, by contrast, asserts the totality of the universal without reference to the exception, but with the crucial proviso that the whole is internally incomplete, or non-all. Lacan relies on these two options, what might be called external exception versus internal excess, to posit two forms of enjoyment, so-called 'phallic *jouissance*' versus the '*jouissance* of the Other'. Put in crude terms, and those only narrowly relevant to sex

itself, Lacan's view here helps explain why masculine desire so typically shows a fetishistic focus on the part object, a tendency to immerse itself in the brute details of the sex act, while the feminine subject remains attuned to the larger context and situation, sensitive to a whole milieu of seduction. On the basis of this structural difference, Lacan posits an irresolvable antagonism between the sexes, tersely expressed in his dictum: *il n'y a pas de rapport sexuel*, 'there is no sexual relation'.

The relevance of Lacan's theory of sexuation to the problem of religion can be sensed in a provisional way when we remember that ancient religions conspicuously and almost universally identify the origin and order of the universe with sexual difference and sexed reproduction. The religious mind, it seems, cannot resist conflating cosmos with coitus. But the relevance becomes more interesting when sexuation is related to Christianity. Once again, we hear an echo of Kant's approach to religion in so far as Christianity turns out to be especially, perhaps even uniquely, reflective of the structures of subjectivity. The dichotomy between phallic and feminine *jouissance* as Lacan defines them can be said to centrally inhabit Christianity, not because Christianity is to be linked to one side or the other but rather precisely because of the way that it straddles the two.

The reason is that the figure of Christ is deeply ambiguous. From one point of view, Christ can be taken to stand for the ultimate exception, the perfected human subject by whose measure all others are to be judged. The Christ-exception stabilizes the law––not, to be sure, by the negative measure of transgression but rather by offering a positive exemplar of perfection. This difference, according to which the role of exception in the grounding of the law is inverted from its more familiar form of a violation or transgression of the law's moral demand, is anything but trivial. In fact, it is by means of this inversion that Christianity achieves a universalization of moral failure, establishing a perfect democracy of sin. Compared with the Messiah, we are all sinners. It is also by this means, as Nietzsche recognized, that Christianity can be thought to bring to completion the Judaic invention of bad conscience (Nietzsche 1887/1992: 493–532). Nevertheless, the logical form of the relation between the universal and the exception remains the same as that described by the masculine formula of sexuation. The universal is founded by the exception.

According to this first reading, Christianity repeats a completely traditional assertion of God as furnishing a master signifier. The organization of the Church as a hierarchy of authority that defends doctrinal orthodoxy would seem to follow with perfect consistency. And yet, a second point of view leads us to an opposite conclusion, as Christ can also be taken to instantiate the ultimate denial of all exception. The figure of Jesus inaugurates a new and unprecedented universalism that is based on his identifying himself with the human. What makes this second view distinct from the first is the way it extends the full implications of the revolutionary Christian notion of incarnation by asserting a kind of infinity inhabiting every human being, a radical equality of all persons in the eyes of God that is born of the fact that every human incarnates something of the divine. This new assertion of the universal, hinting as it does at the sort of sublime unity of the divine and the human alluded to by the mystics, is to be aligned with Lacan's feminine logic of the

non-all. Christ becomes what is 'in me more than me', a tincture of the divine. Paradoxically, a kind of essential, constitutive excess is precisely what makes me human.[20]

It is in the context of this discussion that we can make some sense of Lacan's evocation of a God who doesn't know. On the side of Christian doctrine that links Christ with the masculine logic of exception, God remains omniscient, God is fully *sujet supposé savoir*. 'God only knows', as the familiar saying has it. But according to the more radical potential of incarnation in which Christ-the-suffering-God is identified with all who suffer, Christianity is opened to the feminine logic of excess, an Other *jouissance* that is not knowable. The big Other becomes no longer identifiable with a subject supposed to know, but is, on the contrary, found to be internally incomplete, inconsistent, inhabited by an irradicable vacuity or gap. From this point of view, the Christian divinity ceases to be immutable or impassible. It is for this reason that Lacan's twentieth seminar continually flirts with the notion of an unknowing God, a suffering and kenotic being whose majesty consists precisely in a sublime form of weakness. Lacan thus probes a radical reassessment of divinity arrived at on the basis of psychoanalytic premises, provoking us with the conclusion that 'one can no longer hate God if he himself knows nothing' (1998: 98). As Slavoj Žižek has argued, this assertion of a God who does not know can be seen as a necessary correlate of identifying God with love.

> Only an imperfect, lacking being loves: we love because we do *not* know everything. On the other hand, even if we were to know everything, love would still be higher than complete knowledge. Perhaps the true achievement of Christianity is to elevate a loving (imperfect) Being to the place of God
>
> (Žižek 2002: 61).

The Abyssal Ground of Belief

Early in our discussion of the Lacanian Thing we noted the inadequacy of interpreting it in too Kantian a fashion. It now remains to complete the necessary distancing from Kant and draw out the consequences for our topic. Doing so means asking again about the fuller meaning of the notion of *das Ding*. What order of reality are we dealing with here?

It is immediately tempting to understand the Thing as a substantial entity of some sort. Given Lacan's framing of it in terms of the *jouissance* of the Other, for example, we might pose the Thing in more familiar Freudian terms, taking it as a kind of vortex of excessive libido, a volatile reservoir of surplus energy in the Other that threatens to wash over the subject and overwhelm it.[21] But any such interpretation would miss an absolutely key point. The Thing is in no way a substance, nor even an indeterminate quantum of energy. Nor is it an object. The Thing is the site of a fundamental blind spot, the subject's primitive acknowledgement of a zone of something unknown. The spectre of the Thing results from a kind of primitive projection which essentially says 'there is something here,

Subscriber: University of Otago; date: 17 December 2018

I know not what'. The Thing is thus a reality for the subject alone. It marks the space of an open question. As Freud said of it, the establishment of such a space of the question is what allows the human being to learn to cognize. As such, it performs an absolutely fundamental function, something that might be fruitfully compared to what Heidegger called the disclosive opening of human *Dasein*. It is not itself a being but is rather the lighted clearing in which beings come to presence.

The act by which this site of something unknown is posited cannot easily be named. Indeed, to call it 'posited' already implies too conscious and deliberate an exertion of the part of the subject. But even more difficult is to determine what exactly is posited. The Thing is less an object than a locus, a *topos*. *Das Ding* is not an object, nor is it a fantasy. It is rather the very frame of fantasy, the point of departure or even the cause of fantasy. The Thing is the open question to which fantasy attempts to pose an answer. As such, the Thing is the decentred void that forms the nucleus of the unconscious. One thinks of the 'other scene' that Freud attributed to the dream. In Lacanian terms, the locus of the Thing is 'extimate', something of the most intimate concern to the subject, but located outside, in the Other.

It also remains to clarify the chronology of the Thing. If it is not an object neither is it any sort of prehistorical givenness, some primal point of departure that kicks everything else into motion. Its 'being' is rather an effect of the signifier. The Thing arises with the signifier and is nothing but its excessive dimension. All of which is to say that the Thing is a kernel of the real in the Lacanian sense. And if the real and the symbolic can never be coincident, neither are they separable. They arise together. The symbolic, the real, the subject and the Other constitute an irreducibly four-fold structure. It is Lacan's answer to Heidegger's four-fold of earth and sky, mortals and divinities, and it bears the same consequence of asserting the inescapability of time.[22] This irreducible quadratic dimensionality of the four terms implies that all four are ineluctably caught up in contingent and historical specificity. Their conjunction is historicity itself.

To grasp the full implications of this precision of our discussion is to appreciate the degree to which Lacan's notion is finally less Kantian than Hegelian. Just as Hegel rejected Kant's appeal to a noumenal realm wholly beyond phenomena, Lacan insists that the real is not some impenetrable substance wholly beyond the symbolic but rather arises only from gaps, inconsistencies, and failures internal to the symbolic itself.[23] And we can now make clear how Lacan's assessment of religion parallels that of Hegel, who sharply divides the Christian legacy. On the one hand is the way that Christian incarnation, the Pauline 'scandal for the Jews and folly for the Greeks', opens up an abyssal short-circuiting of the mortal and the divine, the finite and the infinite, even to the point of evoking the death of God. On this side, Christianity marks the truly revolutionary moment in the history of human freedom. On the other is the 'positive', dogmatic aspect of Christianity, its elaboration of doctrinal and theological orthodoxies, which occasioned Hegel's critique in some of his early writings, a critique that resonates with the Pauline dictum that 'the letter killeth but the spirit giveth life' (Hegel 1795/1979).

A Lacanian perspective shows with special clarity why this division is virtually inevitable. It is rooted in the bifold function of the signifier, divided as it is between meaning (corresponding to the linkage of the signifier to a particular signified) and an indeterminate potentiality-for-meaning (the open space of a yet-to-be-determined reference, the space of the Thing). At many points, Lacan emphasizes the first theme, as when, in his interview on 'Triumph of Religion', he characterizes religion as endless spewing out of meaning. The frenzied production of meaning is both what makes religion most seductively attractive and also what makes it so inimical to psychoanalysis. 'Humanity will be cured of psychoanalysis', he says, 'by drowning the symptom in meaning, in religious meaning naturally, people will manage to repress it' (Lacan 2013: 67). But at other times, Lacan appreciates the way in which the religious impulse touches upon the real, a feature that he associates, as we have done here, with the psychic reality of the Thing. As Lacan said of it, 'Freud left us with the problem of a gap once again at the level of *das Ding*, which is that of religious men and mystics' (1992: 100).

In this way, Lacan reveals Christianity as the religion most spectacularly divided between positivity and mystery, dogma and *deinos*.[24] The Christian Church ferociously enforces doctrinal fidelity and fanatically persecutes heresy precisely because the Christian message so radically opens towards something abyssal.[25] The Christian God at once embodies the monstrosity of the Thing, the primary locus of the real, and also the symbolic big Other that seeks to defend against it. Yet the final twist consists in Lacan's refusal to allow any simple opposition between the spirit and the letter. Indeed, the keynote of his entire outlook is to insist on the primacy of the letter, along with all of the *Sturm und Drang* it brings in its train. Without the agency of the signifier, there is no spirit. As he puts it in one of his *Écrits*: 'Of course, as it is said, the letter kills while the spirit gives life. I don't disagree […] but I also ask how the spirit could live without the letter' (Lacan 2006: 423).

References

Balmés, F. (2007). *Dieu, le sexe et la verité*. Paris: Erès.

De Beauvoir, S. (1989). *The Second Sex*. New York: Vintage Books.

Boothby, R. (2001). *Freud As Philosopher: Metapsychology After Lacan*. New York: Routledge.

Chiesa, L. (2016). *The Not-Two: Logic and God in Lacan*. Cambridge, MA/London: MIT Press.

Copjec, J. (2015). 'Sex and the Euthanasia of Reason'. In *Read My Desire: Lacan Against the Historicists* (pp. 201–236). London/Brooklyn, NY: Verso.

Daviot, P. (2006). *Jacques Lacan et le sentiment religieux*. Paris: Erès.

Depoortere, F. (2007). 'The End of God's Transcendence? On Incarnation in the work of Slavoj Žižek'. *Modern Theology 23*(4) October: 497–523.

Freud, S. (1886–99). *Pre-Psycho-Analytic Publications and Unpublished Drafts*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 1. London: Hogarth Press.

Freud, S. (1909). *Two Case Studies*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 10. London: Hogarth Press.

Hegel, G. W. F. (1795/1979). *Early Theological Writings*. Philadelphia, PA: University of Pennsylvania Press.

Heidegger, M. (1962). *Being and Time*. New York: Harper and Row.

Lacan, J. (1968). 'Seminar XVI: From an Other to the other', translated for private use from the unpublished French transcript of the seminar by Cormac Gallagher.

Lacan, J. (1981). *The Four Fundamental Concepts of Psycho-analysis*. New York: W. W. Norton.

Lacan, J. (1992). *The Seminar of Jacques Lacan, Book VII, The Ethics of Psychoanalysis*. New York: W. W. Norton.

Lacan, J. (1998). *The Seminar of Jacques Lacan, Book XX, Encore: On Feminine Sexuality, The Limits of Love and Knowledge*. New York: W. W. Norton.

Lacan, J. (2006). *Écrits*. New York/London: W. W. Norton.

Lacan, J. (2013). *The Triumph of Religion*. Cambridge/Malden, MA: Polity Press.

Lacan, J. (2014). *The Seminar of Jacques Lacan, Book X: Anxiety*. Cambridg/Malden, MA: Polity Press.

Lacan, J. (2017). *The Seminar of Jacques Lacan, Book VIII, Transference*. Cambridge/Malden, MA: Polity Press.

Leupin, A. (2004). *Lacan Today: Psychoanalysis, Science, Religion*. New York: Other Press.

Nietzsche, F. (1887/1992). 'The Genealogy of Morals, Second Essay "Bad Conscience"'. In *Basic Writings of Nietzsche* (pp. 493–532). New York: Modern Library.

Pound, M. (2007). *Theology, Psychoanalysis, and Trauma*. London: SCM Press.

Richardson, W. (1997). '"Like Straw": Religion and Psychoanalysis'. *The Letter: Lacanian Perspectives on Psychoanalysis 11*: 1–15.

Žižek, S. (2000). *The Fragile Absolute—or, Why is the Christian Legacy Worth Fighting For?* London/New York: Verso.

Žižek, S. (2001). *On Belief*. London/New York: Routledge.

Žižek, S. (2002). 'The Real of Sexual Difference'. In S. Barnard and B. Fink (eds), *Reading Seminar XXL Lacan's Major Work on Love, Knowledge, and Feminine Sexuality* (pp. 57–75). Albany, NY: State University of New York Press.

Žižek, S. (2003). *The Puppet and the Dwarf: The Perverse Core of Christianity*. Cambridge, MA/London: MIT Press.

Žižek, S. (2005). 'Neighbors and Other Monsters: A Plea for Ethical Violence'. In S. Žižek, E. Santner, and K. Reinhard, *The Neighbor: Three Inquiries in Political Theology* (pp. 134–190). Chicago, IL: University of Chicago Press.

Žižek, S. (2006). *How to Read Lacan*. New York: W. W. Norton.

Žižek, S. and Gunjevic, B. (2012). *God in Pain: Inversions of Apocalypse*. New York: Seven Stories Press.

Žižek, S. and Milbank, J. (2009). *The Monstrosity of Christ: Paradox or Dialectic?* Cambridge, MA/London: MIT Press.

Notes:

(1.) The past decade has seen a flurry of commentary on the implications of Lacanian theory for rethinking religion. Among the more notable are Pierre Daviot (2006), *Jacques Lacan et le sentiment religieux*; François Balmes (2007), *Dieu, le sexe et la verité*; Marcus Pound (2007), *Theology, Psychoanalysis, and Trauma*; and Lorenzo Chiesa (2016), *The Not-Two: Logic and God in Lacan*. But no one has done more to demonstrate how Lacan reopens the field of the religious from a psychoanalytic point of view than Slavoj Žižek. Žižek treats questions of religion and God in many of his books, perhaps most directly in *The Fragile Absolute—or, Why is the Christian Legacy Worth Fighting For?* (2000), *On Belief* (2001), *The Puppet and the Dwarf: The Perverse Core of Christianity* (2003), *The Monstrosity of Christ: Paradox or Dialectic?* (with John Milbank 2009), and *God in Pain: Inversions of Apocalypse* (with Boris Gunjevic 2012).

(2.) [Eds. See Cottingham, this volume, and Blass, this volume, for more detailed discussions.]

(3.) [Eds. See Eagle, this volume, §'Falsifiability of psychoanalytic propositions', for a discussion of interpretations of the Oedipus complex in relation to empirical testing.]

(4.) 'God is real' states the title of the final chapter of Leupin (2004). For other treatments of God as an embodiment of the real, see Depoortere (2007: 497–523) and Richardson (1997: 1–15).

(5.) For a clear, freewheeling, and highly suggestive introduction to the Lacanian concept of the big Other, see Žižek (2006).

(6.) [Eds. For a very different account of the need for meaningfulness in relation to God, see Cottingham, this volume, §'Freud's critique of religion'.]

(7.) My thesis, which cannot be developed adequately in this chapter, is that while explicit references to the Thing tend to disappear after the seventh seminar, the underlying notion is not abandoned but rather replaced by its representative, the so-called *objet petit a*. When, late in his career, Lacan referred to the *objet a* as his greatest single contribution to psychoanalysis, we should also discern an implicit reference to the concept of the Thing.

(8.) I capitalize 'Other' here and will continue to do so throughout this chapter, but the choice is an awkward one in so far as 'Other' must do double duty between the concepts of the little and big Others. In fact, Lacan himself alternates in his capitalization of *Autre* throughout his work without any apparent consistency. The most logical thing would seem to be using the lower case for the little Other and the upper case for the big Other. But then again, even the little Other of the fellow human being sometimes deserves the emphasis lent by the capitalization, precisely because, when its Thingly dimension is taken into account, the fellow human being becomes something totally different than the impression of ordinary experience leads us to conclude. It is to recognize this point that I will retain the capitalization even of the 'little Other'.

(9.) Linking the primal root of anxiety with the mother herself makes new sense of Freud's observation that the German word '*heimlich*', which by virtue of its reference to 'home' (*Heim*) indicates what is safest and most familiar, but may also, according to its secondary meaning as what remains 'concealed, kept from sight, so that others do not get to know of or about it' function as an equivalent for what is '*unheimlich*, or 'uncanny'. Cf. Freud, (1955: 224).

(10.) 'In the body there is always, by virtue of this engagement in the signifying dialectic, something that is separated off, something that is sacrificed, something inert, and this something is the pound of flesh.' (Lacan 2014: 219).

(11.) 'The most decisive moment in the anxiety at issue, the anxiety of weaning, is not so much when the breast falls short of the subject's need, it's rather that the infant yields the breast to which he is appended as a portion of himself.' (Lacan 2014: 313).

(12.) 'Here the *a* stands in for the subject––it's a stand-in in the position of precedent. The primordial, mythical subject, posited at the outset as having to be constituted in the signifying confrontation, can never be grasped by us, and for good reason, because the *a* preceded it and it has to re-emerge secondarily, beyond its vanishing, marked by this initial substitution. The function of the yieldable object as a piece that can be primordially separated off conveys something of the body's identity, antecedent to the body itself with respect to the constitution of the subject' (Lacan 2014: 314).

(13.) The quotation 'hell is other people' famously belongs to Jean-Paul Sartre. By citing it here I don't mean to conflate Lacan's outlook with that of Sartre. The two are importantly different in ways we cannot expound upon here. Nevertheless, there is a significant, if only partial, overlap with Sartre when it comes to Lacan's treatment of the neighbour-Thing of the fellow human being.

(14.) The precise phrasing of 'the word as the murder of the thing', indeed most of Lacan's acquaintance with Hegel, was made known to him via the highly influential lectures of Alexandre Kojève.

(15.) [Eds. See also Macdonald, this volume.]

(16.) Cf. Lacan's remark: 'That is clearly the essence of law––to divide up, distribute, or reattribute everything that counts as *jouissance*' (1998: 3).

(17.) I have elsewhere put forward an account of sacrifice from a Lacanian point of view. See Boothby (2001: 175–89).

(18.) [Eds. For alternative accounts of the roots of religion in subjectivity, see Cottingham, this volume, and Blass, this volume.]

(19.) Lacan elaborates the formulae of sexuation most extensively in *Seminar XX: Encore*. For especially useful explications of his theory on this point, see Copjec (2015: 201–36), and also Žižek (2002: 57–75).

(20.) Cf. the final chapter of Lacan (1981): 'In You More Than You'.

(21.) We might note in passing that posing the Thing as a mass of energy makes it comparable to the Freudian id, and also invites a link with Lacan's 1955 essay on 'The Freudian Thing' (2006: 334–63) in which the id, *le ça*, is said to speak: *ça parle*.

(22.) Cf. Lacan's remark: 'I have already asked the question here as to what the critical conceivable minimum is for a signifying scale, if the register of the signifier is to begin to organize itself. There cannot be a two without a three, and that, I think, must certainly include a four, the quadripartite, the *Geviert*, to which Heidegger refers somewhere' (2002: 65–6).

(23.) Linking Lacan with Hegel is a tricky business iinsofar as Lacan himself retained a persistently critical view of Hegel as a philosopher of metaphysical closure. Here again, credit needs to be given to Slavoj Žižek, whose clarification of Hegel's thought, convincingly offering an almost complete inversion of the dominant interpretation, has made it possible to recognize the deep conjunction between Hegel's philosophy and Lacan's retheorization of psychoanalysis.

(24.) One wonders whether it is not for reason of Lacan's recognition of the particularly intensity of this double function in the case of Christianity that he claims that 'There is *one* true religion and that is the Christian religion' (Lacan 2013: 66).

(25.) This abyssal dimension is what Žižek and Milbank (2009) has called 'the monstrosity of Christ'.

Richard Boothby

Department of Phlosophy, Loyola University Maryland

Oxford Handbooks Online

Ethics: Introduction ⓐ

Michael Lacewing and Richard G.T. Gipps

The Oxford Handbook of Philosophy and Psychoanalysis
Edited by Richard G.T. Gipps and Michael Lacewing

Print Publication Date: Feb 2019   Subject: Philosophy, Philosophy of Mind, Moral Philosophy
Online Publication Date: Nov 2018   DOI: 10.1093/oxfordhb/9780198789703.013.36

Abstract and Keywords

This introduction provides an overview of the three chapters in this section, which explores central issues in ethics in the context of psychoanalysis, including the nature of virtue, the ground of normativity, moral development, the relation between reason and passion, naturalism and moral motivation. One such issue concerns Sigmund Freud's theory of the superego, which is said to undermine the 'authority' of morality. The first chapter argues that the superego represses conscience, and that our 'moral-psychological difficulties' can be understood only in light of repressed love. The second chapter examines the place of psychoanalysis in the relationship between virtue and mental health, and between vice and mental dysfunction. The third chapter discusses the idea of an 'evolved development niche' to address object relations and their role in moral development.

Keywords: ethics, psychoanalysis, virtue, reason, Sigmund Freud, superego, morality, conscience, mental health, evolved development niche

Michael Lacewing and Richard G. T. Gipps

UNDOUBTEDLY the best-known contribution of psychoanalysis to ethics is Freud's theory of the superego. Many—including Freud at times—have understood his theory to undermine the 'authority' of morality. On this understanding, far from being the voice of pure reason, as Kant would have it, or a God-given sense of what is right and wrong, our conscience as superego is the distillation of the views and values of those around us when we were very young children, distortingly filtered through our own infantile concerns, instincts, and emotions. Furthermore, morality, in the form of the superego, can give rise to various forms of mental distress—narcissism, moral masochism, obsessiveness, ascetism, and a split between 'reason' and 'passion'. Consequently, a reduction in the strength of the superego is a frequent aim of psychoanalysis.

Ethics: Introduction

There are a variety of responses to this picture. The first is to follow the line of thought rather uncritically, and adopt some form of moral scepticism or nihilism. But this tends to be motivated by an oversimple understanding both of morality and of Freud's theory, and so while it has been taken up in popular culture, it is rarely developed by psychoanalysts and philosophers. A second response is to reject Freud's theory of the superego wholesale. But while many of the details have come under attack, e.g. the claim that the superego is the result of resolving the Oedipus complex, those attracted by Freud's naturalism still need some account of our acquisition of moral values. A third response, then, is to develop Freud's theory to find in the resources of psychodynamic psychology a rational alternative (Deigh 1996; Scheffler 1992; Velleman 2006). A fourth approach is to distinguish 'bad superego morality', based on fear and aggression, from 'good ego ideal morality', based on love (Dilman 1984; Erikson 1964; Wollheim 1984). As a variant of this, one might set aside superego theory and use other resources in psychodynamic psychology to remodel our account of the development of our moral sense (Cottingham 1998: ch. 4; Nussbaum 2004: ch. 4).

In his chapter, Backström rejects all of these in favour of a new approach again. He argues that it is a serious mistake to equate conscience and the superego, arguing instead that the superego represses conscience. The foundation of morality is love, understood as an openness to the other, and conscience is the call of love. Backström's approach is to (p. 592) offer a phenomenological analysis of the meanings inherent in our emotional lives to show that our 'moral-psychological difficulties' can be understood only in light of repressed love. We struggle to remain open in love towards others, and the many ways we deny, refuse, and distort this openness not only provide the material for the drama of human relationships, including 'essentially misrecognized' emotions such as envy and jealousy, but also cause many theoretical misunderstandings, from Freud's structural model of id–ego–superego to various accounts of the nature of love to the attempt to ground morality in reason rather than openness. While Backström's account is Freudian in understanding moral psychology in terms of love and repression, it is highly critical of Freud's inability to understand these terms and their relationship adequately.

As argued in our 'Introduction' to this Handbook, questions at the intersection of psychoanalysis and ethics go to the very heart of both disciplines, raising issues about the purpose of each, their self-understanding, and the status of their theoretical models. As Harcourt (2015: 599–600) has commented, there is a 'long tradition' in moral psychology stretching back to Socrates that enquires into 'the relationship between human nature, the good life for man, and human goodness'. Not only has psychoanalysis much to contribute to this enquiry; the enquiry in turn raises challenging issues for how psychoanalysis should understand itself (as a science, as a humanism, as an ethical relationship?) and its aim (psychic health, self-knowledge, goodness?). In his chapter, Harcourt surveys the place of psychoanalysis in this three-way discussion 'between the kinds of creatures we are, what it is to be good or excellent of our human kind, and what constitutes a good life for us'. There are those who argue that psychoanalysis does not aim at nor concern itself with human excellences or the good life, but, by contrast, with psychic health. However, there are others who reject the contrast, because psychic health

should be understood—as Plato and Aristotle thought—in terms of human excellences. Harcourt argues that psychoanalysts who take the former route unknowingly tend to smuggle in human excellences in their characterization of psychic health (Freud 1924; Hartmann 1960). If, then, psychoanalysis addresses the question of human excellences, or again, the question of how to live a good human life, the further question arises whether, on the psychoanalytic account, this is something that is necessarily or recognizably 'morally good'. We are here in the territory of the relationship between virtue and mental health, and its converse, the relationship between vice and mental dysfunction. If there are strong connections here, then psychoanalysis may lend support to the tradition of 'virtue ethics' in philosophy (Erikson 1964; Fromm 1947; Lacewing 2014; Swanton 2011). Harcourt concedes that *some* virtues—those of courage, truthfulness, and the capacity to relate to others as separate beings—may be supported by psychoanalytic accounts of mental health, but argues that otherwise this claim is over-optimistic. What follows? One possibility, which Harcourt is sceptical about but leaves open, is that psychic health, though not a sufficient condition for virtues, is a necessary condition for them.

In the background of Harcourt's discussion is the question of what a naturalistic moral psychology might look like (Lear 2004, 2014). A great deal of empirical psychological work has been done recently on the first of Harcourt's three relata, viz. what kind of beings are we, and its implications for moral psychology. In her chapter, Narvaez begins (p. 593) by arguing that, in studying ourselves as we are now, we have frequently failed to fix an appropriate 'baseline' of comparison by which to understand the results. Seeking to provide a wider empirical context for understanding object relations and their role in moral development, she develops and defends the idea of an 'evolved development niche', the characteristic environment for human babies over the vast expanse of time between our evolution as a species and the changes of the last 10,000 years. Drawing on studies in neuroscience, developmental psychology, cognitive psychology, and anthropology, Narvaez argues that our 'species-typical' upbringing is significantly different from today's culturally typical—but species atypical—parenting practices. Our current practices frequently result in a disordered emotional foundation to the self giving rise to moral psychological problems, a focus on protecting the self against threats rather than an open engagement with others. Her arguments offer the possibility of an empirical framework that supports Backström's phenomenological analysis and takes up Harcourt's question of the connection between virtue and mental health.

Our three chapters thus survey central issues in ethics, including the nature of virtue, the ground of normativity, moral development, the relation between reason and passion, naturalism, moral motivation, and much more. Many other topics in ethics, not covered in any depth in this section, have received rich contributions from psychoanalytic theory. Examples include analysis of moral emotions, such as guilt and shame (Carveth 2013; Harcourt 2007; Wollheim 1999: ch. 3); responsibility (Orange 2014); the aim of human action (Lear 2000); moral epistemology (Lacewing 2005, 2015); the nature of empathy (Richmond 2004); the difference between real virtues and their simulacra (A. Freud 1936:

ch. 10; Fromm 1947); and a great many topics in practical ethics, including, of course, the ethics of confidential relationships.

References

Carveth, D. (2013). *The Still Small Voice: Psychoanalytic Reflections on Guilt and Conscience*. London: Karnac.

Cottingham, J. (1998). *Philosophy and the Good Life*. Cambridge: Cambridge University Press.

Deigh, J. (1996). *The Sources of Moral Agency*. Cambridge: Cambridge University Press.

Dilman, I. (1984). *Freud and the Mind*. Oxford: Basil Blackwell.

Erikson, E. (1964). *Insight and Responsibility*. London: Faber.

Freud, A. (1936). *The Ego and the Mechanisms of Defence*. London: Hogarth Press.

Freud, S. (1924). 'The loss of reality in neurosis and psychosis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 19 (pp. 183–190). London: Hogarth Press/Institute of Psychoanalysis.

Fromm, E. (1947). *Man for Himself*. Greenwich, CT: Fawcett.

Harcourt, E. (2007). 'Guilt, shame, and the "psychology of love" '. In L. Braddock and M. Lacewing (eds), *The Academic Face of Psychoanalysis*. London: Routledge.

Harcourt, E. (2015). 'The place of psychoanalysis in the history of ethics'. *Journal of Moral Philosophy 12*(5): 598–618.

Hartmann, H. (1960). *Psychoanalysis and Moral Values*. New York: International Universities Press. (p. 594)

Lacewing, M. (2005). 'Emotional self-awareness and ethical deliberation'. *Ratio 18*: 65–81.

Lacewing, M. (2014). 'Emotions and the virtues of self-understanding'. In S. Roeser and C. Todd (eds), *Emotion and Value* (pp. 199–211). Oxford: Oxford University Press.

Lacewing, M. (2015). 'Emotion, perception and the self in moral epistemology'. *dialectica 69*(3): 335–355.

Lear, J. (2000). *Happiness, Death, and the Remainder of Life*. Cambridge, MA: Harvard University Press.

Lear, J. (2004). 'Psychoanalysis and the idea of a moral psychology: Memorial to Bernard Williams' philosophy'. *Inquiry 47*: 515–522.

Lear, J. (2014). 'Mourning and moral psychology'. *Psychoanalytic Psychology 31*(4): 470–481.

Nussbaum, M. (2004). *Upheavals of Thought*. Oxford: Oxford University Press.

Orange, D. M. (2014). 'What kind of ethics?: Loewald on responsibility and atonement'. *Psychoanalytic Psychology 31*(4): 560–569.

Richmond, S. (2004). 'Being in others: Empathy from a psychoanalytical perspective'. *European Journal of Philosophy 12*(2): 244–264.

Scheffler, S. (1992). 'Naturalism, psychoanalysis, and moral motivation'. In J. Hopkins and A. Savile (eds), *Psychoanalysis, Mind and Art* (pp. 86–109). Oxford: Blackwell.

Swanton, C. (2011). 'Nietzsche and the virtues of mature egoism'. In S. May (ed.), *Nietzsche's On the Genealogy of Morality: A Critical Guide* (pp. 285–308). Cambridge: Cambridge University Press.

Velleman, D. (2006). 'A rational superego'. In *Self to Self* (pp. 129–155). Cambridge: Cambridge University Press.

Wollheim, R. (1984). *The Thread of Life*. New Haven, CT: Yale University Press.

Wollheim, R. (1999). *On the Emotions*. New Haven, CT: Yale University Press.

Michael Lacewing

Dr Michael Lacewing is a former Vice-Principal Academic and Reader in Philosophy at Heythrop College, London, an Honorary Reader in Research Department of Clinical, Educational and Health Psychology at University College, London, and a teacher of philosophy and theology at Christ's Hospital School, Sussex. He has published widely in philosophy of psychoanalysis, metaethics and moral psychology, alongside writing textbooks for A level philosophy and training in Philosophy for Children (P4C).

Richard G.T. Gipps

Dr Richard Gipps is a clinical psychologist in private psychotherapy practice in Oxford, UK, and an associate of the Philosophy Faculty at the University of Oxford. He convenes the Philosophy Special Interest Group of the Institute of Psychoanalysis, the Oxford Interdisciplinary Seminars in Psychoanalysis, and the Making the Unconscious Conscious seminar series. His research interests lie in psychoanalysis, psychosis, existential phenomenology, and Wittgenstein.

Oxford Handbooks Online

Abstract and Keywords

Freud's account of morality is distinctive, and right, in focusing on unconscious, emotionalized conflict, and specifically on the repression of love as the centre of moral life. However, Freud misunderstands love in drive terms and confuses conscience with the superego. Conscience is actually an immediate moral understanding, an interpersonal openness that the moral normativity of collectivity (values, ideals, etc.) represses. Thus, conscience is the repressed unconscious of the superego, and 'morality' not one thing, but a living contradiction. This chapter details how bad conscience differs from superego guilt, how destructive emotions (e.g. jealousy) are in themselves moralized repressions of love, and how Freud's officially amoral, drive-based accounts of the Oedipus complex and the installation of the superego break down, but can be understood if reconceptualized in the terms proposed here. The chapter elucidates the concrete sense in which openness and love can be conceived as the very heart of moral understanding.

Keywords: morality, love, openness, repression, conscience, superego, guilt, emotions, Oedipus complex, collectivity

Joel Backström

(p. 595) The Superego and its Other

FREUD accepts the traditional view of morality as a set of norms/ideals that constrain or redirect our supposedly natural egocentric inclinations. Freud's version, where norms are enforced by the 'superego' against sexual and aggressive 'drives' housed in 'the id' poses a sceptical, naturalizing challenge to the idealizing–intellectualizing tendency of traditional ethics, which sees moral norms not as mere social impositions, but as authored by Reason and/or God. Against this, Freud presents conscience as 'an embodiment, first of parental criticism, and subsequently of that of society', its 'voice' the

inner echo of parental voices and the clamour of 'public opinion' (Freud 1914b: 96). Naturalistic views of morality are standard today. Freud is distinctive in focusing on how moral norms come to *concern* the child at all, and consequently on the intimate dynamics of family life, seen in terms of destructive inner conflicts which morality both regulates and perpetuates. Morality is a socially necessary but cruel system of internalized repression rather than, even ideally, a satisfying development of capacities for moral reasoning and virtuous character. Its core workings are unconscious, its rigid demands not the product of reason but a primitive way of managing libidinal deadlocks, and its 'reward': madness for some and for the majority, general discontent and widely ramified cultural pathologies born of an obscure, mostly unconscious sense of guilt.[1]

(p. 596) A bleak picture! A common response is to think that Freud exaggerated; at best, his description might be accurate for individual pathology and particularly vicious versions of collective morality—after all, he lived in repressive Victorian times (cp. Lear 2015: 190–7). For many psychoanalysts, making 'rigid', 'punitive', 'archaic' superegos more 'realistic', 'humane', and 'mature' is a central goal of therapy,[2] and philosophers revise Freud's theory to allow for a 'rational superego', i.e. for moral reasoning rather than mere blind obedience in our response to socially sanctioned norms.[3] Theorists make the picture of moral development friendlier and (they believe) more realistic by stressing our natural inclination to sociability ('attachment') against Freud's idea of the amorally egocentric child, the pre-Oedipal relationship to the nurturing caregiver(s) over the hostile Oedipal rivalry that transfixed Freud, and the role in moral education of positive ideals and encouragements to identification (the aspirational 'ego ideal', even a 'loving superego') over threats and prohibitions.[4]

I propose something different. A broadly 'Freudian' focus on unconscious, emotionalized conflict is urgently needed, lest philosophical ethics remain stuck in the back-and-forth between equally naive rationalisms and sentimentalisms. Freud's account of ethics is seriously confused, however, his notion of childish amorality being a case in point (more on that later). But making it more 'reasonable' or otherwise 'positive' obscures the crucial insight he points to concerning the destructiveness of everyday collective morality, where we pressure others and ourselves into submitting to depersonalizing decrees about how one 'must' think and behave. This isn't merely irrational or too harsh, but a deeply destructive falsification of life. This comes into full view, however, only in contrast to the wholly other, interpersonal moral understanding that I will call *conscience*, which collective moralism, i.e. the superego, attempts to *repress*. Morality is a thing much more fundamental *and* more terrible than either idealists or cynics ever suspect, because it isn't really one thing at all, but a living contradiction, comprised of our impossible attempts to deny and destroy the openness between us which conscience calls us to welcome. Freud never says anything like this. Paradoxically, he sees the destructiveness of collective moralism, but not what it is ultimately destructive of. The basic defect of his account is that he replaces conscience with its false double, the superego, and this is intimately connected with his confused view of love. His core idea that human life is a drama centred on *love and its repression* is right, but he misunderstands its meaning.

Others, too, have criticized Freud for not distinguishing 'conscience' from 'superego',[5] and many have insisted, using different terms, that there are 'two quite distinct types of (p. 597) morality' (Reddish 2014: 1).[6] Indeed, something like that idea motivates the pervasive wish to 'improve' the superego, moving towards 'mature' morality. But the decisive question is how one conceives this contrast. Is the superego a primitive form of conscience, say, or a repressive reaction-formation against it, as I propose? Once the basic structure of motivations constituting the superego is clearly seen, I think 'improving' it will seem as appealing as a bigger cell in prison when compared to freedom. Hopefully, my account also indicates why morality has, basically, nothing to do with being *reasonable*, and why there's no need for an 'independent faculty of normative judgement, located in the ego' that would 'subject conscience to judgement'.[7]

In my view, 'conscience' refers to one's personal moral understanding of one's relation to the other person, as it appears when one is tempted to go against it by somehow closing oneself to the other. If, seeing your need, I simply help, without reservations or ulterior motives, we wouldn't speak of 'conscience', yet the bad conscience that pricks me if I let you down expresses the same responsive openness that in the other case made me simply help. This openness is, very concretely, our 'moral sense'; the sense one person makes of another and of their encounter because she *senses* how it is with the other, responds to him in an immediate way that precedes inferences and reasoning. Thus, your sadness moves me in compassion; your gentleness warms me and makes me glad; your indifference chills me. These aren't mere passive reactions, but active forms of understanding and being with the other; in openness I enter your gentleness and sadness, but I'm closed out by your indifference. This openness is always already and inescapably there, insofar as any sense we make of what transpires between us presupposes and expresses it—if only, often, in the form of our destructively repressing it. This is simply how we relate to and understand each other.[8] This interpersonal understanding is itself *moral*, rather than merely psychological, insofar as feelings and judgements which are 'moral' in a narrower, often moralistic, sense (e.g. shame and guilt) remain dependent on it for their moral significance. If someone claimed to understand he'd done wrong in brutally beating someone, yet in his response to his victim showed no compassionate understanding of *her*, felt no chill as he thought of what he'd done, his supposed 'understanding' would be morally empty.[9]

However, this inseparability of conscience and interpersonal understanding is obscured precisely insofar as conscience is confused with the guilt- and shame-inducing superego, which happens in everyday life before being echoed by theorists like Freud. The slaveholder's sense that he's duty-bound to keep slaves in their place, the guilt and shame (p. 598) he might feel for being 'too soft' on them, certainly don't express the same understanding as his compassion with the grief of the slave mother whose child he just sold down the river. On the contrary, these destructive moralistic feelings, and the values, ideals, and 'moral reasons' that correspond to them on the 'intellectual' side, function to repress the slaveholder's interpersonal moral understanding, hardening him against the slave and against his own heartfelt understanding of her. Thus, in identifying with his

collective belonging, he identifies with his superego and represses his conscience. Conscience is the repressed unconscious of the superego.[10]

This doesn't mean that human beings are naturally good, and only get ruined by an evil society. Society exists only through/as the way individuals relate to each other and themselves, and the evil I do is indeed mine; I make it mine in identifying with 'my station' and its evil duties. That is: I close myself to those I make my victims, but because it's impossible to *simply ignore* others (to simply *stop* one's compassion, etc.), closing oneself necessitates, as its other side, a constellation of collectively enabled, destructive identifications and projections. Thus, I see my victim by turns as dangerous and contemptible, disgusting and attractive, pitiable and having himself to blame, and my view of myself and the collective I identify with is correspondingly ambivalent; e.g. I see 'us' as superior to 'them' *and* as defenceless against their aggression. A fruitful ethics needs to understand precisely the intra- and interpersonal emotional dynamics at stake in such conflicting, self-obfuscating identifications/projections. And this means understanding how, starting in the nursery, love is repressed and relationships are depersonalized.

Jealous Oedipus: Repressing Love through Moralized Emotion

The focus on life as structured and deformed by repression—i.e. by the unconscious created by our desperate need to keep ourselves unaware of what troubles us—is, Freud says, 'the corner-stone on which the whole structure of psycho-analysis rests' (Freud 1914a: 16).[11] And Freud always relates repression to love; repression is the manifestation of an 'incapacity for loving' (Freud 1904: 267), and 'every psycho-analytic treatment is an attempt at liberating repressed love' (Freud 1907: 90). In other words, what we suffer from is love trouble. Alas, this trouble infects Freud's various statements about (p. 599) the matter, too, turning 'love' (*Liebe*) into a symptom and agent of confusion in his text, a sign whose meaning is perpetually changing and slipping away. Unsurprisingly, theorizations of love are (de)formed by one's personal wishful and fearful fantasies about love: cynical, romantic, sentimental, etc. You may wear the theorist's hat, but the head and heart under it are your own.

In Freud's earlier texts, 'love' is supposed to mean 'libido', that is, sexuality understood as a 'drive' composed of 'component drives', originating in somatic 'erotogenic zones' and variously sublimated and displaced (Freud 1905). Then, in rehauling his drive theory and introducing the controversial 'death drive', Freud also gives 'love' a new, broader signification (Freud 1920; cp. Lear 2003: 157ff.). Instead of reducing love to sexuality, sexuality now becomes an aspect of love understood as a general life force, Eros, striving 'to establish ever greater unities … to bind together', while the death drive aims 'to undo connections and so to destroy things' (Freud 1930: 122; 1940: 148). Freud's later drive conception is notoriously speculative and obscure; we'll return to its problems. His earlier reduction of love to libido was just as problematic, however; in fact, mere theoretical

fantasy. Had Freud actually thought about love as reductively as he pretended to, he couldn't have made any sense of his clinical material.

The best illustration is Freud's depiction of the Oedipus complex. The breakdown of his libidinal account of it reveals the general impossibility of conceiving morality as, most basically, the imposition of norms on amoral 'nature'. Freud officially holds that morality enters the inner world of the previously amoral child only through the installation of the superego, which ends the Oedipal phase, making the superego 'the direct heir of the Oedipus complex' (Freud 1924a: 167; cp. Freud 1923: 28–39; 1924b). The Oedipal constellation itself is supposedly a purely libidinal, amorally sexual-emotional affair (including its aggressive aspects). But Freud's actual descriptions reveal something altogether different; a drama of terrible tensions and difficulties between family members, driven by jealous rivalry which arises because those involved both love *and* mistrust each other. The Oedipal child, Freud says, proceeds as though the parents, primarily the parent of the opposite sex, should by rights be devoted to it exclusively, and regards anything contradicting this as proof of 'unfaithfulness' (Freud 1910: 171; cp. Freud 1917: 332–4; 1924b; 1933a: 118–135, etc.). Pleasure, which Freud claims sexuality, and so the whole libidinal drama, is ultimately about, isn't at stake, but love; love *repressively misrepresented* as preference and possession, where the value of pleasures given is inextricable from their being perceived as tokens of privileged status. The Oedipal drama, then, isn't about sexuality but about love trouble—or, as one might also say: human sexuality, that trouble-ridden thing, is created by the presence of love and love trouble (Backström 2014a).

The critical point isn't limited to sexuality, but concerns any attempt to understand interpersonal dynamics in drive terms, that is, taking the individual's private/anonymous urges and needs as primary, with the other person entering only derivatively, as a contingent provider of satisfactions. Drives, Freud says, become attached to people —'objects'—'only in consequence of [the object] being peculiarly suited to make satisfaction possible' (Freud 1915: 122; cp. Freud 1905: 147–9). But his own descriptions show that the (p. 600) Oedipal drama turns precisely around one's troubled relation to the other person; that relationship determines *what* one finds satisfying or frustrating (cp. Freud 1924b: 176). Unlike Fairbairn (1994), I'm not asserting the primacy of the 'object' over the 'drive'; those are strictly correlative concepts, the 'object' being what the 'drive' focuses on. The primary thing is the actual relationship to *the other person*. Drive objects are formed as repressive responses to love trouble in that relationship and function, like the values, ideals, etc. to which they're typically tied, to repress conscience (which calls one to love). Love trouble means closing oneself to the other, reducing her, in fantasy, to a certain *kind* of person. In one's irritation or excitement, say, one feels 'driven' towards, or away from, this irritating or exciting 'object' to which one's own difficulty in being open with the other reduces her. By contrast, there's nothing *driven* (compulsive) or objectifying in love.[12]

Love trouble is *moral* trouble because it concerns the relation to the other person. Jealousy, envy, hatred, and the related emotions and fantasies that constitute the Oedipal drama are in themselves morally defined responses. They aren't morally *good*, but they manifest, however confusedly and wilfully, an accusing sense of entitlement, of one's rightful position vis-à-vis others; Freud speaks of 'romances of revenge and exoneration' (Masson 1985: 318). Prior to any external superego judgement, this accusatory moralism is constitutive of what the jealous (etc.) child *feels*. Jealousy and the other Oedipal attitudes are morally charged because they are *demands raised in the name of love* (not *out of* love).

The Oedipal boy doesn't just want the presence of mother and isn't resentful towards father simply for being the cause of her absence. He is jealous because he sees that mother cares about father, loves him. In his jealous eyes, *that* is the betrayal. His demand for devotion is essentially triangular and comparative; he demands that mother *exclude* father in favour of him. He might be indifferent or irritated with mother as long as they're alone, but once father enters, he suddenly cannot be without her. Thus, the rival's presence, self-deceptively presented and felt in jealousy as simply a hateful threat, is actually what makes the supposedly loved person 'lovable' to the jealous. The jealous demand for devotion doesn't indicate strength of love, and even where there's a strong wish to be with the other also in the rival's absence, jealousy reveals a *distrust* of the love between oneself and the other. If two people love wholeheartedly, with no secret reservations and anxieties between them, a third person cannot become a 'rival' any more than buyers can exist in the absence of sellers.

However, while the jealous demand for devotion isn't an expression of wholehearted love, it is *related* to love. A child who felt no love, but was simply the biological discharge-seeking machine Freud claims children originally are (Freud 1933a: 62), would simply use people for his own ends. He wouldn't feel jealous, for he wouldn't care what happened between his parents, unless it directly impinged on his private interests. He (p. 601) couldn't feel excluded from their love because he wouldn't perceive it as *love* but only as, say, 'an opportunity to snatch a cookie while they're occupied with each other'. Actual children aren't like that; like adults, they love, but also find it difficult to love, i.e. are susceptible to jealousy and myriad other emotional responses in which love is played on and perverted by fearful, destructive egocentricity. Egocentricity isn't amorality, but a repression of love in which one indulges, as no amoral creature could, feelings of rage and self-pity over fantasized betrayals by others whom one claims to love, but destructively attacks in one's self-pitying rage, thereby repressing the love one does *also* feel.[13]

The jealous boy presents himself, both to mother and to himself, as having been abandoned by her and accuses her, as Freud says, of 'lack of love' (Freud 1933a: 122). The accusation is fraudulent; even if mother really acted lovelessly, the jealous child's accusing attitude is itself loveless. Compare the forlorn, heartbreaking expression of a child who genuinely feels abandoned by a parent, with the simultaneously self-pitying and tyrannical demeanour of a jealous child, furiously protesting if the parent shows interest

in others. No difference could be greater! With a heartfelt sense of abandonment comes anguish, but no *demand*. The abandoned child's look is heartbreaking precisely because it doesn't try to *force* the parent to stay; it simply longs for the parent, suffers defencelessly from her callousness. The jealous child appeals to this very defencelessness, makes a *show* of it, and parents often play along; 'Those sad, innocent puppy eyes, I just couldn't say no!' Alas, that isn't love speaking, and there's nothing innocent about this game of mutual moralistic sentimental manipulation. The jealous demand makes it clear that unless you yield, you'll be sorry, but the threat is *moralized*. It doesn't say: 'If you leave, I'll break your favourite vase', but rather: 'Look how *hurt* I would be by your leaving; think of how sorry, how *guilty* that would make you feel!'

The jealous child doesn't only accuse the parent, trying to make her feel guilty; he also feels guilty himself. Because of the repressed love in jealousy and other responses, they have a self-accusing aspect; they are essentially *guilty* feelings. The Oedipal boy doesn't only want, egocentrically, to have mother's total devotion; he also loves her, and so wants to see her happy. And he loves and cares for his father, too, however intensely he may *also* hate him as a rival. Indeed, hatred is itself a repressed form of love; its perversely insatiable concern with the hated other, with whom one *claims* to want nothing to do, arises as one's loving concern for the other, which cannot be simply abolished, is repressively turned into a destructive persecuting and being persecuted by him.[14] The presence of love means that, if the jealous boy gets mother to neglect father in preference to him, his triumph will be mixed with guilt for having made both father and mother sad by forcing himself between them and ruining the moment of love between them (to repeat: if he didn't perceive this love, he wouldn't feel *jealous*). The destruction he wrought makes the boy sad, gives him bad conscience. If he allowed (p. 602) himself to simply perceive what his bad conscience shows him, to feel it without repressing it, he would feel no guilt-ridden jealousy or triumph, but wholehearted love in the form of a longing for forgiveness and reunion with both his parents, where *no one* would be excluded. If this sounds like a naively idyllic notion, consider a situation where a jealous boy's 'death-wish' is seemingly 'granted', and father actually goes away for good; this will certainly not make the boy happy, but deeply unhappy and guilt-ridden, and he will wish for nothing more than to have father back.

Jealousy, like other responses, is attended by bad conscience, but one can persist in jealousy only by repressing this bad conscience, this longing for lovingly reuniting with those one jealously turned against. The repression leaves one feeling guilty, but not, consciously, for the *jealousy*, which one presents, *in* feeling it, as a justified response to the other's treason. Instead, one feels guilty for other things, proxies, and one may not feel overtly guilty at all, but rather one's guilt comes out in self-punishing behaviour, etc. Thus, the 'unconscious sense of guilt' with its symptomatic displacements, which Freud always emphasized (cp. Johnston 2013: 88–101), is itself a repression of love, that is, of conscience.

The crucial point is that this 'primal repression' of love's bad conscience is inherent in jealousy and other responses; it is part of having/feeling them at all. This is why triumph over a rival and other 'good' feelings engendered by destructive, self-centred schemes and fantasies never feel simply good, but always also terrible, intense and even euphoric as they may be. The repression of love, the self-obfuscation or splitting-of-oneself, is part of the feeling itself. Here, then, repression doesn't work according to Freud's schema, where '–instinctual impulses' undergo repression 'following the commands of [the] super-ego' (Freud 1924c: 150) because they 'come into conflict with the subject's cultural and ethical ideas'—which makes 'the formation of an ideal … the conditioning factor of repression' (Freud 1914b: 93–4; 1925: 29; 1895: 268–9). While one may indeed repress awareness of one's jealousy because one has internalized a moral prohibition against jealousy (an ideal of being a non-jealous person), the primary repression happens already *in* the jealous response itself: to feel jealous *is* to repress the very love in the name of which one raises one's jealous demands. Hence, jealousy and other reactions can be characterized as *essentially misfelt/misrecognized responses*. One cannot be clear about them *in* having them; having them means misrepresenting to oneself the character of the relation to the other person they manifest.[15] Thus, jealousy involves misrepresenting oneself as loving the person one blackmails with one's jealous demands and confusing love, which is between 'you' and 'me', with preference, i.e. rejection of a third party (without this confusion, there'd be no jealousy). Egocentric/destructive responses repress both the actual lovelessness of one's own response *and* the love that nonetheless also underlies it. Thus, in one's hatred one feels not loveless, but justified by the *other's* supposed lovelessness ('After what she did to me … '), and one refuses to acknowledge the love one still also feels for her.

(p. 603) In my view, our moral difficulties generally are structured by responses; they involve a repression of love and conscience and are constitutively self-misrepresentations. This also means that love and destructiveness aren't two equiprimordial 'drives', as Freud imagines (Freud 1930: 122). Rather, the myriad forms of our destructiveness are repressive responses to love, parasitic on the very relation to the other they desperately, impossibly, want to destroy. Reik's claim that 'Love is in its essential nature an emotional reaction-formation to envy, possessiveness, and hostility' (2002: 66), gets it exactly the wrong way round.

The Superego's Empty Threat

Jealousy is not a one man show. Like our 'inner' life generally, it's a response within an evolving relationship to others, and their responses to us change the field, the sense, of our further responses. If others didn't confuse the jealous child's manipulations with love but clearly saw them for what they are, they'd lose their effectiveness, appearing merely sad. The confusion is self-deceptive; manipulation, unlike simple deception, plays on the manipulated person's wish to go along with the manipulation.[16] Parents are taken in by their manipulative children not because they're stupid, but because they themselves are playing the same manipulative game. Freud tends to depict the Oedipal constellation

from the child's point of view, but parents who feel rejected by a child in favour of the other parent are as prone as the child to jealousy, and if the jealous boy feels 'special' in having mother privilege him, mother feels 'special' in seeing the boy insistently demanding her for himself. Neither of them wants to give up the narcissistic satisfactions of their self-deceptive game. And similar games are everywhere; think of how easily telling someone something trivial gains an aura of importance if presented as a *secret*; lowering my voice, I say 'I've told no one else', and you feel flattered to be let in on this, in contrast to all those excluded. The Mother–Child–Father triangle doesn't exist in a vacuum, and siblings, relatives, friends, and others are drawn into and actively play along in the games of mutual jealousy, envy, preferences, sympathies, and other forms of egocentrically, distrustfully comparing attitudes, where one is judged by others and lowered or raised in their affection and esteem relative to others, just as one in turn judges them.

Let's now consider how the 'superego'—the culturally variable constellations of more narrowly moral values and norms with their corresponding patterns of guilt and shame that children internalize—fits into this web of deformed relationships. Freud claims that 'the destruction of the Oedipus complex is brought about by the threat of castration'; faced with threatened punishment, the child eventually gives up its parent(s) as object(s) of jealous sexual fantasies of possessing mother/father and instead identifies with them (p. 604) in their role as authorities prohibiting this very possession; 'The authority of … the parents is introjected into the ego, and there … forms the nucleus of the super-ego' (Freud 1924b: 176–7; cp. Freud 1923: 28–39). But fear of threatened punishment (Freud understands castration literally; we needn't) cannot create moral understanding, and 'introjection' changes nothing here. Introjecting a frightening, punitive adult creates a frightening, punitive 'internal object' that appears in the child's anxieties and nightmares, but it doesn't make it feel the least *guilty* (cp. Harcourt 2007; Jones 1966). Guilt feelings arise only where punishment is administered in the name of love, or, rather, only then can aggressive reprisal be experienced as *punishment*. Insofar as adults could simply mistreat children, with nothing like affection between them, children would fight for their survival without moral scruples, as though confronted by some terrible machine. Freud himself finally admits this when, after trying unsuccessfully to explain the origin of the superego only in terms of aggression turned back on the child's self (1930: 123–30), he acknowledges the ineliminable part 'played by love in the origin of conscience [i.e. the superego]' (Freud 1930: 132). As he says, only 'the experience of being loved' can turn aggression into guilt (Freud 1930: 130 fn).[17]

The superego's basic motive force is, Freud says, fear of 'loss of the parents' love' (Freud 1930: 124–5). Its 'pangs of conscience' are really stabs of this internalized fear, while in heeding superego demands one 'expects to be rewarded by receiving more love from it', and 'consciousness of deserving this love is felt … as pride' (Freud 1939: 117). This description is simultaneously both accurate and confused, because it characterizes the destructive dynamics at issue in the same repressively confused terms used by those who submit others and themselves to it.[18] The question we should ask is how *love* could be withdrawn, as opposed to repressed? Certainly, favoured status can be withdrawn, out of

some motive or capriciously. But love isn't favouritism or conditionality. Anyone who says 'I'll love you *if* you do *xyz*' isn't speaking the language of love. In turning my supposed 'love' into a weapon of extortion, I only show that I *don't* love you. In other words, Freud hangs everything on an empty threat; the threat of taking away what the threat itself reveals isn't there, or is only there in repressed form, as *refused*.

Alas, Freud's conception seems plausible because we constantly threaten each other in this way, to terrible effect. This essentially empty threat is more powerful than all 'real' threats combined. Hurt looks and other manipulative strategies of moralized, emotionalized blackmail imply that if you want to be 'loved', you have to make yourself 'lovable'; change yourself to fit *my* preferences, wishes, needs. However, insofar as the situation is moralized, I don't present these as simply mine, but as about you meeting shared expectations and values; '*No one* accepts a person who acts like you!' In this moralistic depersonalization one enlists the support of others against the person one threatens, making her feel isolated and powerless. But depersonalization isn't primarily driven by (p. 605) strategic concerns; the basic point is that in making egocentric demands one refuses to be open, to address the other simply as oneself, 'I' to 'you', and this *means* that one starts speaking in general terms, about what 'one'—perhaps as 'a person in my position, after the kind of thing you did'—can and cannot do, think, understand, etc. This depersonalized mode of address doesn't change one's basically egocentric, manipulative orientation, but manifests it. To repeat: in rejecting you because you don't meet my expectations, whether shared with 'everybody' or with no one, I show that I refuse to love you, not that you're 'unlovable'.

Does loving someone then mean uncritically accepting abuse and evil from them? No. Love rejects evil unconditionally, but it doesn't reject *the other* even when she does evil. Rather, evil prevents love's realization insofar as, in doing evil, the *evil-doer* refuses to give or receive love. In speaking to someone's conscience in love's sense, one calls her back to herself and to the one(s) she deserted. By contrast with superego voices, there's no moralistic demand or threat, and no argument, just a simple appeal: 'Don't you see what you're doing?' It's an appeal without power, but precisely this freedom from—*not* mere lack of—power makes it in a way stronger than any power. Think, again, of the appeal in the forlorn child's face; in not attempting to manipulate and force the other, but leaving them free, it opens them to the question of their own freedom and responsibility. This breaks their heart open—or drives them mad, if they refuse the opening.

In succumbing to parental threats of 'loss of love', the child feels ashamed or guilty; the empty threat becomes 'real', effective, through those emotional responses. They contrast with two other possibilities. A child might react amorally, obeying without shame or guilt, simply to avoid the external consequences of parental disapproval, or it might be 'unable' to repress its love and so reacts to parental shaming/guilt-loading with sadness and anguish, perhaps anger. The child who, instead, represses its love in succumbing to shame/guilt, turns on its own spontaneous expressions, anxiously trying to remake itself in the image of 'the lovable child' projected by the parents' disapproval of the real one; thus arises the ego ideal. The tragedy is that while the child longs for love, by falsifying

itself it gets only *approval*. Adam Smith thought we have an 'original desire to please, and … original aversion to offend our brethren' (1976: 116), and that this is the foundation of morality, but the desire to please isn't original; it's what our longing for love gets perverted into insofar as we turn it down. 'Society' doesn't do this to us; rather, society in its destructive aspects is constituted and perpetuated by our collectively doing it to ourselves. The jealous child, too, with its hurt looks and angry protests, threatens the parent with 'withdrawing its love', claiming the right to punish the faithless parent for 'withdrawing her love first'. Indeed, every move in our collective threatening game is presented as justified by the lovelessness of the other. Thus, the punitive mother presents herself as *punishing* her child for his naughtiness; that is, for acting selfishly, not caring about her or others.

The paradigmatic superego responses shame and guilt suffuse the threatening game. These responses are in some ways quite different. Most obviously, shame focuses on one's weakness/ineptness in not measuring up to some ideal, i.e. to the expectations of those whose rejection one fears, guilt on the harm one has supposedly done them; not (p. 606) one's impotence but the destructive use made of one's power makes one feel guilty (thus, if A made B cry, A feels guilty for her meanness, her power to hurt, while B feels ashamed for his weakness). But the differences between shame and guilt arise from a deeper similarity. In both cases one fears social rejection and (so) loss of self-esteem; one feels an intolerably *bad* person. This is obvious in shame, and it means that, even where my shame results from my mistreating you in what I regard as a shameful way, my thoughts aren't primarily with you in your suffering; rather, I'm preoccupied with the misfortune I've brought on myself (*I* have lost face). This narcissism in the *form* of one's moral concern makes shame morality a terrible thing even where the particular values ('contents') it honours are in some ways desirable. A murderer who feels *ashamed* of his deed only perpetuates the blindness to his victim, and to himself, it revealed.[19]

It might seem, though, that a murderer *should* definitely feel guilty. We shouldn't confuse guilt feelings with bad conscience, however; the former are actually a repression of the latter, and share shame's self-loathing self-centredness. Obviously, a repentant wrongdoer knows he's guilty in the plain sense of having done someone wrong. But the more he opens up to the person he wronged, i.e. the more he allows himself to feel love for her— that love is *what* he feels in his bad conscience, what it calls him to—the less will he be tormented by guilt feelings, which are feelings focusing, like shame, on one's own (supposed) badness. The primary expression of bad conscience is the longing to find those one wronged, ask their forgiveness, and be reconciled with them. Here, one isn't dominated by the egocentric worry over others withdrawing their 'love', i.e. their favour. Instead, there's a pained realization that, by hurting, ignoring, or manipulating the other, one has *oneself* failed in love for *them*. This realization isn't intellectual, but manifests one's desire to open up to the other in love again. The sense of having closed oneself to the other and being still estranged from them is painful; the more one loves, the clearer and deeper is the pain. But since one now wholeheartedly seeks reconciliation and sees the other and oneself in one's relation to them in the light of love, one's consciousness of

estrangement—unlike guilt and shame, but like grief or longing for an absent friend—doesn't feel *simply* bad (cp. Nykänen 2014).[20]

The guilt- and shame-inducing superego's function is, as Freud says, *hostile surveillance*; one feels watched and condemned by others (Freud 1930: 136; cp. Freud 1933a: 59). The repression and displacement of conscience by the superego consists precisely, as Nykänen (2014) demonstrates, in this reversal whereby, instead of actively opening oneself in love to the person one wronged, one assumes the passive position of 'one judged by others'. These others needn't be real people, although at first they were that for the child; they may be imagined, idealized others, representatives of impersonal moral values. (p. 607) Philosophers debate whether the moral judgement one feels subject to is 'autonomous' (impersonal) or 'heteronomous' (tied to an audience of specific others), but from love's perspective that difference is irrelevant; however the judging 'one' is further specified, the self-objectifying, judgemental perspective contrasts with the bad conscience of love, where one doesn't feel observed and condemned, but opened to and united with the other in love. Furthermore, even in its most instinctive and infantile forms, comparing oneself with more or less 'lovable' (favoured) others involves a depersonalizing self-objectification; the Oedipal boy already implicitly wonders 'What does he (the rival) have that I lack, that makes mother prefer him?' Apparently, some 'thing', some quality or ingredient, *makes* one 'lovable' (valuable); the ego ideal is built around the core fantasy of acquiring this 'thing', this value, thereby *forcing* others to 'love' (value) one. 'Autonomy' is the fantasy of *making oneself valuable* by fashioning oneself into an object that others 'must' value by a kind of impersonal, ideal necessity.

Some may find linking bad conscience to forgiveness cheap; saying 'I'm sorry' might make one feel good, but what counts is making reparations and bettering one's ways. The objection is misguided. Asking for forgiveness isn't just *saying* 'I'm sorry' and perhaps shedding some guilty tears. When spoken from the heart, 'Please forgive me' is the unadorned expression of the longing to undo the distance and alienation between oneself and the other caused by oneself. One makes no claims, doesn't pressure the other to forgive or pretend that they are *obliged* to; one simply expresses the same longing of love expressed in the forlorn child's 'Please don't go'. And that longing will also be expressed in endeavours to change entrenched ways of acting, to repair what was broken, etc. Where the longing is absent, or rather repressed, as when guilt feelings dominate and there's no forgiveness sought, reparation, too, will acquire a different, morally problematic character. Then, one will try to appease the guilt one feels for having *avoided* seeking forgiveness, compensating for it by 'doing good'. Think, e.g., of a woman forever 'paying' in the coin of kindness and self-abnegation to her husband for an affair she never dared confess and ask forgiveness for, because truthful confrontation with him frightens her.

Love: All Freedom, No Licence

The main point I've been labouring is that our moral psychological difficulties, in pathological and everyday forms, can be understood only in the light of the good possibility they repress, namely the wholeheartedness of love to which conscience calls us. That possibility is what we have difficulties *with*. Thus, both the pressure we often feel to lie and the terribleness of lying relate to the truthful, wholehearted contact the lie evades and represses. Love is the good that the various forms of human destructiveness, moralized or not, are destructive of.

But, it will be objected, can't people do terrible, as well as wonderful, things out of love? And isn't the 'first desideratum' in any account of love therefore 'to recognize that love comes in good and bad varieties', and to decide 'which ideal sorts love into good and (p. 608) bad' (Harcourt 2017: 39)? Freud certainly thought so, and until we see the misconception underlying this view, it remains unclear how conscience in the sense of love could provide genuine moral understanding. The crucial point is that so-called bad 'forms of love' aren't really forms of love, but attitudes that masquerade as love while actually repressing it. As I've stressed, terrible things are done *in the name* of love. No word is as abused as 'love'. To tell the real thing from the counterfeits one doesn't need an 'ideal', however, but the real thing itself. Someone who loves wholeheartedly won't be taken in by false protestations of 'love', while one who doesn't will use even the best ideals to rationalize the distrust, defensiveness, and destructiveness which manifest their own lovelessness. 'If someone has no ideals', Frankfurt says, 'there is nothing that he cannot bring himself to do' (1999: 114). But it's precisely in the name of ideals, often loveless ideals of 'love', that the worst atrocities are committed.

It isn't merely that some ideal representations of love—those tied up with oppressive gender roles, say—misrepresent love, but that any *representation* of love is itself a *repression* of love. Any ideal of love comes between the lovers, obstructing and perverting the openness between them. Loving someone means approaching them without preconceptions and demands, conscious or unconscious, regarding how one 'should' treat them or be treated by them, or what 'kind' of person oneself and they are, whereas ideals are precisely pieces of 'legislation' about these things. To love is to listen and speak, to touch and move, freely, unrestrictedly, guided only, and fully, by love's wholehearted desire to know and be known by the other, in every sense of word, and this includes the endless work of undoing the clogs of distrust and misunderstanding that always threaten to build up between us. Here, even the best ideals and the most positive expectations and 'narratives' are, ultimately, only hindrances.[21]

Certainly, without moral ideals, norms, and prohibitions, social life, including intimate relationships, would descend into barbarism. But that's because there's so much lovelessness (jealousy, envy, etc.) between people, not because *love* must be regulated. Moral regulation by norms/ideals is like the police: necessary, given how heartless and irresponsible we often are, but also a symptom of the love trouble they only precariously control. Clearly, insofar as things are *good* between people, no police surveillance is

needed, and if it's still insisted on, that itself becomes a problem. The idea that, if left free, love *itself* might lead to abuse, is confused; abuse is excluded by the very dynamics of love's desire. If you love someone you also desire sexually, for instance, there's no need for you *also* to (say) 'respect' them to ensure your desire doesn't lead to violations, for when sexual arousal is lived in love, *what* you desire—that the other turn to you in the openness, the unguarded desire and joy, in which you turn to them—excludes every violation. What you're aroused by is the prospect of sexual communion, and if the other doesn't want that, it kills your arousal. Or rather, your arousal ceases where it would become, through the other's withdrawal, *merely your own, your private thing*. Thus, insofar as my desire to caress you expresses love, I won't go on if you don't want me to, because (p. 609) there will now *be* no place for me to go. You were the 'place', and you've withdrawn. By contrast, if my caress is loveless, i.e. if love is only present in repressed form, my response to your unwillingness may vary. If I wish to ingratiate myself with you, I'll anxiously stop and apologize, perhaps; if I feel sentimental and cuddly, or just horny, I might go on *because I want to* so much; indeed, in my egocentric arousal I may not even notice that you don't like what I'm doing. The more I give myself over to my sentimental or horny fantasizing, the more I ignore *you*, on whose person I stage my fantasy (Backström 2014a).

Wholehearted love thus subverts the standard conception of the relation between desire and morality, which sees a person's desires, needs, etc. as essentially egocentric and therefore amoral. On that conception, what I desire and experience as 'good for me' is compatible with the other's good only contingently, and so concern for the other must be expressed through external moral regulation (through taboos, reasoning, or in some other way). Even apparently 'positive' desires like the desire to relieve distress rooted in empathic identification, need moral regulating because they're still essentially egocentric and capricious. I might feel empathic towards A and B, but not C; my empathy for A might disregard his actual predicament (perhaps I wrongly see him as like myself), etc. This privatizing conception of desire, shared by rationalists and sentimentalists, doesn't fit love's desire, however, which is *itself* the welcoming opening of oneself to the other. That's why, far from needing external moral regulation, it is the very movement of morality, of conscience. Conscience can reveal the moral destructiveness of both my private inclinations and the shared moral norms of my society not because it's some 'higher' form of subjectivity, whatever that would mean, but because it *is* my openness to the other person, the one my private or collectively sanctioned aims tempted me to violate—as I now clearly see, in the light of love that makes her presence real to me.[22] That light excludes all licentiousness/irresponsibility. Moral regulations, by contrast, always leave some room for it, within the limits they establish; thus, my obligation to respect your wealth is matched by your right to use it as you please, e.g. keeping it all while your neighbours starve.

Love is the longing for the other to be herself and oneself to be oneself, together. Thus, love is a longing for *reality*, 'a spirit of truth … which will not … have anything to do with … falsehood' (Weil 1978: 242); the very opposite of the blindly idealizing credulity Freud and many others present it as (Freud 1905: 150, 295–6; 1921: 111–16). That view leads

Freud to suggest that love itself can drive repression, as when a husband blinds himself to hateful aspects of his wife's person; in Freud's view, 'precisely the intensity of his love … would not allow his hatred … to remain conscious' (Freud 1909: 180–1). But it isn't out of love that the husband is hostile and represses his hostility; it's because he isn't prepared to love his wife wholeheartedly. He cannot bear to see her and his feelings for her truthfully, because that would mean opening things up between them; he'd have to change, to forgive, and ask for forgiveness. Instead, he represses the problem, which (p. 610) means, correlatively, that he fabricates a false, idealized picture or narrative about their relationship. He may admit they have problems, but he'll misrepresent what they're about. Refusal of love is compatible with brutal frankness in acknowledging shortcomings in oneself and the other; one can admit to any kind of 'truth' *except* the truth that one refuses to love (not cannot, but refuses to). And it isn't this truth as such that frightens one, but the prospect of love itself, of opening up to the other in forgiveness.

Love, then, is what is repressed, never what represses. When you find something amiss in the other's way of relating to you, or to someone else, the question of conscience in love is whether your own reaction expresses love—in which case what you react to is the other's closing herself to love—or is rather a way in which *you* close yourself to her, while projecting your guilt over this onto her, presenting her as 'harassing' or 'judging' you, say.[23] And you know what's what only by opening yourself in love, as conscience calls you to, for opening up *means* letting go of your fearful hostility with its repressive/projective games, thus fully and lucidly feeling what is going on between you. There is then no 'moral question' left, only the longing to be and stay open. This is what the dissolution of moral difficulties looks like from the perspective of love (cp. Nykänen 2015). This dissolution doesn't tell you what to *do* in the situation, for instance how best to help someone in need. But the *moral* aspect of the difficulty consisted in your not really wanting to help. Perhaps the other disgusted you or provoked your envy, and so you instinctively made up justifications for not helping; 'He doesn't deserve help', 'He'll manage', etc. Opening up to the other dissolves the disgust/envy, and along with it the falsely moralized pseudo-problem of whether he 'deserves'/needs help, and you're finally able to look at the problem as a genuinely practical one of how best to help him.

A common objection to equating conscience with love is that two people loving each other isn't enough and that lovers may treat 'outsiders' indifferently—even viciously, if they feel their love threatened, so that 'in the blindness of love, remorselessness is carried to the pitch of crime', as Freud says (1921: 113). But this, too, is misconceived. Unlike favouritism or devotion which indeed tend towards remorselessness against 'outsiders'— as do collective us/them identifications generally, however wide in scope—love is non-excluding, even if, or, rather, precisely because, it's strictly personal. No extra moral demand to 'extend' the love one feels for one person to others is needed, because the very form of love's desire excludes all destructiveness, not just against one particular person ('the beloved'). Loving even one person wholeheartedly means opening *oneself*, i.e. shedding one's fearfully insistent egocentric demands, and this changes the spirit in

which one approaches each person one meets; one now sees *each* other *as* other (Backström 2007: 229–316).

And, turning things around: whence derives the objector's certainty that everyone *should* be treated well and equally? How does one know, say, that the slaveholder's attitude to his slaves is evil? Isn't that because love shows one this in conscience, as one considers what it would mean to treat a loved human being as the slaveholder treats his ⁽ᵖ. ⁶¹¹⁾ slave? Repressing this understanding means repressing one's sense of the slaveholder's *callousness*, and this allows his attitude to appear justified, even perfectly reasonable—as it did to Aristotle, whose blindness to the evil of 'natural slavery' (cp. Cambiano 1987) wasn't due to any weakness in his rational powers, but to identification with his slaveholding culture. Aristotle's superego forbade him to see its evil. The social pressure internalized in the superego determines the limits allowed to reasoning rather than the reverse; it is 'rational' only in the sense that, speaking in the superego's depersonalizing voice, we enlist reasoning as readily as emotional reactions for our repressive purposes. Reasoning doesn't tell ideologizing rationalization from moral understanding if one doesn't *want* it to. And if one wants to see the difference, no reasoning is needed. One doesn't perceive callousness as a result of arguments, and if one starts *arguing* the wrongness of slavery, one has already conceded too much, insofar as one admits, in principle, that the slaveholder might be right (if one doesn't admit that, one isn't really reasoning, merely rationalizing).

The 'Impossibility' of Ethics

I have been radically critical of Freud's account of morality. However, my critique is itself Freudian in inspiration insofar as I understand moral difficulties in terms of love trouble and repression. My criticism is that Freud doesn't see *how* love and repression are actually related, and how both can be properly understood only as essentially moral phenomena—just as, conversely, moral life itself can only be understood in terms of love and its repression.

Let me finally illustrate a crucial 'methodological' point about the study of ethics by briefly commenting on Freud's structural model of the psyche. That model doesn't model *the* structure of our minds, as Freud thinks, but is rather itself a symptom of real-life difficulties; a theoretical formulation of the *self-misrepresentation* belonging to destructive moral-existential orientations. The model reifies 'ego', 'id', and 'superego' as independent psychical 'agencies', but they're really only different stances (or 'voices') we assume as we address ourselves and others. It is *I* who speak in an impatiently horny voice ('id'), or in a prurient voice condemning horny desires as 'immoral' ('superego'). However, I thereby become a different kind of 'I' than I would be if I dared to confront the other openly; an 'I' for which the term 'ego' is quite apposite. The egocentric attitude I assume is marked by irresponsible passivity, self-pity, and hostility. Picturing myself as an innocent bystander caught in the middle of a conflict between the 'natural urges' I 'simply find welling up in myself' and the arduous moral demands for renunciation that

'society' imposes, I refuse to take responsibility for either my desire or my moral understanding. Freud's structural model doesn't *account* for this depersonalization and flight from responsibility, from openly addressing the other in one's longing, sorrow, or anger; it merely *echoes* it. The model is the self-obfuscating self-apology of the 'poor ego' who often 'cannot suppress a cry: "Life is not easy!" ' (Freud 1933a: 78).

(p. 612) What this brief discussion illustrates was implicit in all the discussions here; that the difficulty of doing ethics isn't primarily intellectual. The pictures we make of morality, including at the most abstract theoretical levels, are themselves implicated in our love trouble, defended and rejected depending on what wishful and fearful fantasies we've unconsciously invested in them. This means that accounting for morality truthfully 'in theory' implies a kind of 'working-through' of one's own moral-existential difficulties and temptations; a sustained engagement precisely with what one most wants to avoid thinking about. If psychoanalysis is an 'impossible' profession, so is ethics, and essentially for the same reason.

References

Alexander, F. (1925). 'A metapsychological description of the process of cure'. *The International Journal of Psychoanalysi*s *6*: 13–34.

Alford, C. (2006). *Psychology and the Natural Law of Reparation*. Cambridge: Cambridge University Press.

Backström, J. (2007). *The Fear of Openness: An Essay on Friendship and the Roots of Morality*. Åbo: Åbo Akademi University Press.

Backström, J. (2014a). 'Fearful fantasies: Sexuality as a response to love'. *Scandinavian Psychoanalytic Review 37*: 48–59.

Backström, J. (2014b). 'Touchy subjects: Wittgenstein's "Freudianism" and the Moral Dynamics of Repression'. *European Journal of Psychoanalysis 2* (2014), **http://www.journal-psychoanalysis.eu/**.

Backström, J. (2015). 'Of dictators and green-grocers: On the repressive grammar of values-discourse'. *Ethical Perspectives 22*: 39–67.

Backström, J. (2017). 'From nonsense to openness—Wittgenstein on moral sense'. In E. Dain and R. Agam-Segal (eds), *Wittgenstein's Moral Thought*. New York: Routledge.

Backström, J. and Nykänen, H. (2016a). 'Collectivity, evil and the dynamics of moral value'. *Journal of Evaluation in Clinical Practice 22*: 466–476.

Backström, J. and Nykänen, H. (2016b). 'Ethics: the perfect mix? Response to Buetow'. *Journal of Evaluation in Clinical Practice 22*: 479–482.

Barnett, B. (2007). *'You Ought To!' A Psychoanalytic Study of the Superego and Conscience*. London: Karnac.

Bergler, E. (1952). *The Superego: Unconscious Conscience*. New York: Grune & Stratton.

Bieber, I. (1972). 'Morality and Freud's concept of the superego'. In S. M. Post (ed.), *Moral Values and the Superego Concept in Psychoanalysis* (pp. 126–143). New York: International Universities Press.

Boag, S. (2012). *Freudian Repression, the Unconscious, and the Dynamics of Inhibition*. London: Karnac.

Brickman, A. S. (1983). 'Pre-oedipal development of the superego'. *The International Journal of Psychoanalysis 64*: 83–92.

Britton, R. (2003). *Sex, Death, and the Superego: Experiences in Psychoanalysis*. London: Karnac.

Cambiano, G. (1987). 'Aristotle and the anonymous opponents of slavery'. In M. I. Finley (ed.), *Classical Slavery* (pp. 21–41). London: Frank Cass.

Carroll, J. (1985). *Guilt: The Grey Eminence Behind Character, History and culture*. London: Routledge & Kegan Paul. (p. 613)

Carveth, D. (2013). *The Still Small Voice: Psychoanalytic Reflections on Guilt and Conscience*. London: Karnac.

Church, J. (1991). 'Morality and the internalized other'. In J. Neu (ed.), *The Cambridge Companion to Freud* (pp. 209–223). Cambridge: Cambridge University Press.

Cordner, C. (2002). *Ethical Encounter: The Depth of Moral Meaning*. Houndmills: Palgrave.

Cottingham, J. (1998). *Philosophy and the Good Life*. Cambridge: Cambridge University Press.

Deigh, J. (1996). *The Sources of Moral Agency: Essays in Moral Psychology and Freudian Theory* (ch. 4 and ch. 6). Cambridge: Cambridge University Press.

Dilman, I. (2005). *The Self, the Soul and the Psychology of Good and Evil*. Abingdon: Routledge.

Fairbairn, W. R. D. (1994 [1952]). *Psychoanalytic Studies of the Personality*. London: Routledge.

Feldman, K. S. (2006). *Binding Words: Conscience and Rhetoric in Hobbes, Hegel, and Heidegger*. Evanston, IL: Northwestern University Press.

Frankfurt, H. (1999). *Necessity, Volition, and Love*. Cambridge: Cambridge University Press.

Freud, S. (with J. Breuer) (1895). 'Studies on hysteria'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 2 (pp. 3–311). London: Hogarth Press.

Freud, S. (1904). 'On psychotherapy'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 7 (pp. 257–268). London: Hogarth Press.

Freud, S. (1905). 'Three essays on the theory of sexuality'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 7 (pp. 125–245). London: Hogarth Press.

Freud, S. (1907). 'Delusions and dreams in Jensen's *Gradiva*'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 9 (pp. 7–95). London: Hogarth Press.

Freud, S. (1909). 'Notes upon a case of obsessional neurosis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 10 (pp. 153–249). London: Hogarth Press.

Freud, S. (1910). 'A special type of choice of object made by men'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 11 (pp. 165–175). London: Hogarth Press.

Freud, S. (1913). 'Totem and taboo'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 13 (pp. ix–162). London: Hogarth Press.

Freud, S. (1914a). 'On the history of the psycho-analytic movement'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 7–66). London: Hogarth Press.

Freud, S. (1914b). 'On narcissism: an introduction'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 67–102). London: Hogarth Press.

Freud, S. (1915). 'Instincts and their vicissitudes'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 117–140). London: Hogarth Press.

Freud, S. (1917). 'Introductory lectures on psychoanalysis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vols 15 and 16 (pp. 13–496). London: Hogarth Press.

Freud, S. (1920). 'Beyond the pleasure principle'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 18 (pp. 3–64). London: Hogarth Press. (p. 614)

Freud, S. (1921). 'Group psychology and the analysis of the ego'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 18 (pp. 67–143). London: Hogarth Press.

Freud, S. (1923). 'The ego and the id'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 19 (pp. 3–66). London: Hogarth Press.

Freud, S. (1924a). 'The economic problem of masochism'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 19 (pp. 156–170). London: Hogarth Press.

Freud, S. (1924b). 'The dissolution of the Oedipus complex'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 19 (pp. 173–179). London: Hogarth Press.

Freud, S. (1924c). 'Neurosis and psychosis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 19 (pp. 149–153). London: Hogarth Press.

Freud, S. (1925). 'An autobiographical study'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 20 (pp. 7–74). London: Hogarth Press.

Freud, S. (1926). 'Inhibitions, symptoms and anxiety'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 20 (pp. 87–174). London: Hogarth Press.

Freud, S. (1930). 'Civilization and its discontents'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 21 (pp. 59–145). London: Hogarth Press.

Freud, S. (1933a). 'New introductory lectures on psychoanalysis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 22 (pp. 3–182). London: Hogarth Press.

Freud, S. (1933b). 'Why war?' In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 22 (pp. 199–215). London: Hogarth Press.

Freud, S. (1939). 'Moses and monotheism'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 23 (pp. 3–137). London: Hogarth Press.

Freud, S. (1940). 'An outline of psycho-analysis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 23 (pp. 144–207). London: Hogarth Press.

Gaita, R. (2004). *Good and Evil: An Absolute Conception*, 2nd edn. London: Routledge.

Gillman, R. D. (1982). 'Preoedipal and early Oedipal components of the superego'. *Psychoanalytic Study of the Child 37*: 273–281.

Gray, P. (2005). *The Ego and Analysis of Defense*, 2nd edn. Northvale, NJ: Jason Aronson.

Grinberg, L. (1964). 'Two kinds of guilt: their relations with normal and pathological aspects of mourning'. *The International Journal of Psychoanalysis 45*: 366–371.

Grotstein, J. S. (2004). 'Notes on the superego'. *Psychoanalytic Inquiry 24*(2): 257–270.

Harcourt, E. (2007). 'Guilt, shame, and the "psychology of love" '. In L. Braddock and M. Lacewing (eds), *The Academic Face of Psychoanalysis*. London: Routledge.

Harcourt, E. (2015). 'The place of psychoanalysis in the history of ethics'. *Journal of Moral Philosophy 12*(5): 598–618.

Harcourt, E. (2017). 'Attachment, autonomy and the evaluative variety of love'. In E. Kroeker and K. Schaubroeck (eds), *Love, Reason and Morality*. New York: Routledge.

(p. 615)

Holder, A. (1982). 'Preoedipal contributions to the formation of the superego'. *Psychoanalytic Study of the Child 37*: 245–272.

Holmes, J. (2011). 'Superego: An attachment perspective'. *The International Journal of Psychoanalysis 92*(5): 1221–1240.

Jacobsen, E. (1964). *The Self and the Object World*. Madison, CT: International Universities Press.

Johnston, A. (2013). 'Misfelt feelings: unconscious affect between psychoanalysis, neuroscience, and philosophy'. In A. Johnston and C. Malabou (eds), *Self and Emotional Life: Philosophy, Psychoanalysis, and Neuroscience* (pp. 73–224). New York: Columbia University Press.

Jones, D. H. (1966). 'Freud's theory of moral conscience'. *Philosophy 41*: 34–57.

Kernberg, O. (1976). *Object Relations Theory and Clinical Psychoanalysis*. New York: Jason Aronson.

Kesel, M. de (2009). *Eros and Ethics: Reading Jacques Lacan's Seminar VII*. Albany, NY: SUNY Press.

Lacewing, M. (2008). 'What reason can't do'. In N. Athanassoulis and S. Vice (eds), *The Moral Life: Essays in Honour of John Cottingham* (pp. 139–165). Basingstoke: Palgrave Macmillan.

Lacewing, M. (2014). 'Emotions and the virtues of self-understanding'. In S. Roeser and C. Todd (eds), *Emotion and Value* (pp. 199–211). Oxford: Oxford University Press.

Lear, J. (2000). *Happiness, Death, and the Remainder of Life*. Cambridge, MA: Harvard University Press.

Lear, J. (2003). *Therapeutic Action: An Earnest Plea for Irony*. London: Karnac.

Lear, J. (2015). *Freud*, 2nd edn. Abingdon: Routledge.

Levinas, E. (1969). *Totality and Infinity*. Pittsburgh, PA: Duquesne University Press.

Løgstrup, K. E. (1997 [1956]). *The Ethical Demand*, trans. T. Jensen, G. Puckering, and E. Watkins; rev. H. Fink and A. MacIntyre. Notre Dame/London: University of Notre Dame Press.

Masson, J. M. (1985). *The Complete Letters of Sigmund Freud to Wilhelm Fliess 1887–1904*. Cambridge, MA: Harvard University Press.

Milrod, D. (2002). 'The superego'. *The Psychoanalytic Study of the Child 57*: 131–148.

Neill, C. (2011). *Lacanian Ethics and the Assumption of Subjectivity*. Houndmills: Palgrave.

Novick, J. and Novick, K. K. (2004). 'The superego and the two-system model'. *Psychoanalytic Inquiry 24*(2): 232–256.

Nykänen, H. (2002). *The 'I', the 'You' and the Soul: An Ethics of Conscience*. Åbo: Åbo Akademi University Press.

Nykänen, H. (2005). 'Heidegger's conscience'. *Sats 6*: 40–65.

Nykänen, H. (2009). *Samvetet och det dolda—om kärlek och kollektivitet*. (*Conscience and the Hidden: On Love and Collectivity*). Ludvika: Dualis.

Nykänen, H. (2014). 'Conscience and collective pressure'. *Philosophy, Psychiatry, & Psychology 21*: 51–65.

Nykänen, H. (2015). 'Repression and moral reasoning: An outline of a new approach in ethical understanding'. *Sats 16*: 49–66.

Reddish, E. (2014). *The Petrified Ego: A New Theory of Conscience*. London: Karnac.

Reik, T. (2002 [1949]). *Love and Lust: On the Psychoanalysis of Romantic and Sexual Emotions*. New Brunswick: Transaction Publishers.

Reiner, A. (2009). *The Quest for Conscience and the Birth of the Mind*. London: Karnac.

Roth, P. (2001). *The Superego*. Duxford: Icon Books.

Sagan, E. (1988). *Freud, Women, and Morality: The Psychology of Good and Evil*. New York: Basic Books. (p. 616)

Sartre, J.-P. (1965 [1946]). *Anti-Semite and Jew*, trans. G. Becker. New York: Shocken Books.

Schafer, R. (1960). 'The loving and beloved superego in Freud's structural theory'. *Psychoanalytic Study of the Child 15*: 163–188.

Schecter, D. (1979). 'The loving and persecuting superego'. *Contemporary Psychoanalysis 15*: 361–379.

Scheffler, S. (1992a). *Human Morality* (ch. 5). Oxford: Oxford University Press.

Scheffler, S. (1992b). 'Naturalism, psychoanalysis, and moral motivation'. In J. Hopkins and A. Savile (eds), *Psychoanalysis, Mind and Art: Perspectives on Richard Wollheim* (pp. 86–109). Oxford: Blackwell.

Sedlak, V. (2016). 'The psychoanalyst's normal and pathological superegos'. *The International Journal of Psychoanalysis 97*: 1499–1520.

Segal, H. (2002/1979). *Klein*. London: Karnac.

Smith, A. (1976 [1759]). *The Theory of Moral Sentiments*, eds. D. D. Raphael and A. L. Macfie. Oxford: Clarendon Press.

Strachey, J. (1934). 'The nature of the therapeutic action of psycho-analysis'. *The International Journal of Psychoanalysis 15*: 127–159.

Suttie, I. (1963). *The Origins of Love and Hate*. Harmondsworth: Penguin.

Symington, N. (2004). *The Blind Man Sees: Freud's Awakening and Other Essays* (ch. 3 and ch. 6). London: Karnac.

Velleman, D. (2006). *Self to Self: Selected Essays*. Cambridge: Cambridge University Press.

Warner, C. T. (1982). 'Feelings, self-deception, and change'. *Association of Mormon Counselors and Psychotherapists Journal* April 1982: 21–31, 35.

Warner, C. T. (1986). 'Anger and similar delusions'. In R. Harré (ed.), *The Social Construction of Emotions* (pp. 135–166). Oxford: Basil Blackwell.

Weil, S. (1978 [1952]). *The Need for Roots*, trans. A. F. Wills. London: Routledge.

Wollheim, R. (1986). *The Thread of Life*. Cambridge: Cambridge University Press.

Wurmser, L. (2000). *The Power of the Inner Judge*. Northvale, NJ: Jason Aronson.

Žižek, S. (2005 [1994]). *The Metastases of Enjoyment: On Women and Causality*. London: Verso.

Zupančič, A. (2000). *Ethics of the Real: Kant, Lacan*. London: Verso.

Notes:

(^{1}) Three central Freud texts here are (1923, 1930, 1933a: 57–80); see also his (1913, 1914b, 1921, 1924a) etc. Barnett (2007), Church (1991), and Roth (2001) are recent sympathetic readings of Freud on morality. On individual/cultural symptoms of superego guilt, see Bergler (1952), Carroll (1985), Wurmser (2000), and Žižek (2005).

(^{2}) For example, Grotstein (2004), Jacobsen (1964), Kernberg (1976), Strachey (1934), and Sedlak (2016).

(^{3}) Velleman (2006), cp. Deigh (1996) and Scheffler (1992a, 1992b).

(^{4}) For example, Bieber (1972), Brickman (1983), Cottingham (1998), Dilman (2005), Gillman (1982), Harcourt (2015), Holder (1982), Holmes (2011), Lacewing (2008, 2014), Lear (2000, 2015: 211–4), Schafer (1960), and Schecter (1979).

(^{5}) For example, Carveth (2013), Reiner (2009), Sagan (1988), and Symington (2004). Lacanians don't speak of conscience, but they too project an ethics 'not subject to the logic of the superego' (cp. Kesel 2009; Neill 2011; Žižek 2005; Zupančič 2000: 160).

(^{6}) Novick and Novick (2004) distinguish 'closed-system' from 'open-system' superegos; Kleinians 'paranoid-schizoid, persecutory' from 'depressive, reparative' guilt (cp. Alford 2006; Grinberg 1964; Segal 2002), etc.

(^{7}) Velleman (2006: 149) and Britton (2003: 101). Cp. Alexander (1925), Gray (2005), and Milrod (2002).

(^{8}) This isn't an empirical claim but a philosophical remark concerning the space of human intelligibility as such, within which any meaningful empirical testing must move (cp. Backström 2017).

(^{9}) Gaita (2004: 43–63) underlines this point. Levinas (1969) and Løgstrup (1997) also make the concrete relation to the other person basic, whereas most ethics proceeds as though others gained moral significance only in relation to *my* values (cp. Backström 2015; Nykänen 2005).

(^{10}) This general perspective on ethics, focusing on our difficulties with openness to the other and our destructive responses to it, is developed at length in Backström (2007), following Nykänen (2002). Cp. Backström and Nykänen (2016a, 2016b).

(^{11}) By 'repression' (*Verdrängung*), Freud sometimes means a specific defence in contrast to others, but he mostly uses the word generically, as I do, covering an open-ended range of defensive strategies (cp. Boag 2012; Freud 1926: 163–4). Freud's discussions of

repression are inadequate precisely because he overlooks that *what* is ultimately repressed is conscience. As Symington says, 'We cannot do evil ... and know it. We fashion the unconscious so as not to know it' (2004: 70; cp. Backström 2014b; Nykänen 2009).

([12]) Freud's later drive conception might *seem* to mark 'the movement from a closed to an open psyche' (Lear 2003: 166), with Eros conceptualized as a *non*-egocentric/reductive/ destructive drive, but conceiving love in drive terms still focuses on the singular subject; Eros becomes 'the quasi-instinctual force by which ... the psyche grabs hold of meaning to organize itself [... in] ever more diverse communications with itself and with the world at large' (2003: 177–8). Here, the psyche is open to itself and 'its' world, not to the other person—which is what *love* is all about, and why it's the very heart of moral understanding.

([13]) I'm not making empirical claims here (or, I hope, anywhere else in my argument), but clarifying the immanent connections of meaning between various motives and attitudes; connections we typically deny in indulging these very attitudes.

([14]) Cp. Cordner (2002: 147–9), Løgstrup (1997: 32–5), and Sartre (1965).

([15]) Cp. Johnston (2013), Nykänen (2009), and Warner (1986, 1982).

([16]) Selling sugar pills as painkillers is simple deception; taking the pills and feeling the pain subside 'as a result' is self-deceptive self-manipulation; pill merchants who count on this response are engaged in manipulation.

([17]) Cp. Freud 1930: 124–33. Many commentators overlook the centrality of love in Freud's account of superego formation (e.g. Church 1991; Symington 2004), and Freud himself sometimes appears to deny it (e.g. Freud 1933b: 211).

([18]) Even writers critical of Freud's reductive view of love accept the loss-of-love idea; e.g. Suttie (1963: 102) and Harcourt (2007: 143).

([19]) Clearly, people *can* feel ashamed even for murder; confusion is confused, but one can *be* confused. Similarly, philosophers regularly overlook or explicitly deny the narcissism of shame, even praising it as 'the prime moral sentiment of evolved morality, of morality beyond the superego' (Wollheim 1986: 220).

([20]) Some may feel that the 'bad conscience' I contrast with self-centred guiltiness is actually guilt in the *proper* sense of the word. Since I'm not legislating about the use of words, but trying to characterize two different moral orientations, I wouldn't object to this—if my description of the contrast is otherwise accepted.

([21]) Hence, the ubiquitous psychoanalytic talk of 'good objects' and 'good narratives' should be treated with caution.

([22]) Huckleberry Finn is an exemplary case (Backström 2007: 240–8, 338–343). Many philosophers rightly criticize subjectivistic misconceptions of conscience, mistakenly thinking they're criticizing conscience itself (cp. Feldman 2006).

(^{23}) In actual cases, love and loveless motives will typically be mixed on both sides; I don't wish to deny this ambivalence but to understand what creates it, *what* is mixed.

Joel Backström

Department of Philosophy, University of Helsinki, Finland

Oxford Handbooks Online

Human Excellence and Psychic Health in Psychoanalysis ⧉

Edward Harcourt

The Oxford Handbook of Philosophy and Psychoanalysis
Edited by Richard G.T. Gipps and Michael Lacewing

Abstract and Keywords

This chapter begins with a seeming rivalry between two answers to the question 'what are the aims of psychoanalysis?', which seem to situate psychoanalysis differently in relation to ideals of human excellence: 'we are trying to make people good'; and 'goodness is none of our business—we are just trying to make people healthy'. But are these alternatives? If, as Aristotle said, human excellence is psychic health, one aim cannot be achieved without the other. That still leaves a space between human excellence and 'the moral virtues', until it is shown that one cannot be excellent of our human kind without possessing the moral virtues—as modern philosophers assume, though as Nietzsche denied. Attempting to resist these various assimilations, this chapter aims to untangle the complex relations between psychic health as it is conceived in psychoanalysis, the possession of the 'moral virtues', and excellence of our human kind.

Keywords: Freud, psychoanalysis, Aristotle, virtue ethics, psychic health, morality, Hartmann, teleology, character, human development

Edward Harcourt

(p. 617) Introduction

PLATO and Aristotle initiated an investigation into the triangular relations between human nature, distinctively human excellence, and the human good: between the kinds of creatures we are, what it is to be good or excellent of our human kind, and what constitutes a good life for us. In recent philosophy, at least in the Anglophone world, the burden of this investigation has been carried by what is often known for better or worse as 'virtue ethics'.[1] Its central exhibit—with which labourers in the 'virtue ethics' vineyard may agree or disagree—is the ambitious thesis that the life both in which we are distinctively excellent of our kind and that is best *for* us is the life in which we possess

certain presumptively valuable dispositions of character. Though there's room for controversy about exactly which ones make it on to the list (modesty?) and why those that make it on to the list do make it (because these dispositions enable us to meet our impartial moral obligations? because they are necessary for us to lead our characteristic species lives?), if a core consisting of at least justice, fidelity to promises, honesty, courage, and generosity *didn't* make it on to the list, many practitioners of 'virtue ethics' would agree that the ambitious thesis wasn't worth arguing about.

From within the trenches of recent Anglophone philosophy, it can look as if 'virtue ethics' is the *only* place nowadays in which the Platonic–Aristotelian investigation is carried on. But more than one psychoanalyst (Lomas 1999: 3; Lear 2005: 514) has claimed (p. 618) that the question with which patients come to analysis is Socrates' 'how to live' (Plato 1961: 352d)—that is, surely, the 'good life for us' corner of the Platonic–Aristotelian triangle. Again, according to Hans Loewald (1978: 5–6), 'psychoanalytic findings and theory, are … [about] becoming what may properly be called a self, a person'.[2] That sounds like a claim about the proper development of our distinctively human natures, and so—assuming, plausibly, a connection between excellence and proper development (Wall 2013: 342)—a claim about excellence of our human kind. Psychoanalysis[3] has thus been seen, by some at least, as another continuator of the Platonic–Aristotelian investigation. Indeed, it has a fair claim to be a continuator that's more visible to all but the professional philosopher than 'virtue ethics' itself.

To set against this claim, however, others—beginning with Freud—have maintained that psychoanalysis is animated not by an ideal of human excellence, nor of the proper development of our distinctively human natures—however exactly that is to be understood—but by an ideal of psychic health. On this view, psychoanalysis is interested not in people becoming better as people, but only in their becoming less ill. In Freud's words, it is 'a procedure for the medical treatment of neurotic patients' (1963a: 15; cp. 1958: 115).

Moreover, even between those psychoanalytic thinkers most readily associated with the Platonic–Aristotelian investigation and their 'virtue ethics' counterparts, there seems to be a striking difference in respect of which real or apparent human excellences matter.

To avoid confusion, we need to pause here on a point of terminology. To the ear attuned to the 'virtue ethics' literature, there may be no audible difference between 'the virtues' or indeed 'the moral virtues' on the one hand and, on the other, 'human excellences' or 'excellences of character'. Thus the *Stanford Encyclopedia of Philosophy* explains a virtue as 'an excellent trait of character' and continues—seemingly as if no new idea is being introduced—that 'a virtuous person is a *morally* good, excellent or admirable person' (Hursthouse and Pettigrove 2016). However, this assimilation of virtues or excellences to *moral* virtues or *moral* excellences leaves us without a ready label for those traits of character—such as the ability to enjoy oneself, adventurousness, Aristotle's 'proper ambition', or (a contemporary favourite) 'resilience', i.e. the ability to recover from setbacks—which are presumptively good, but which don't seem like a fit for 'morally

good', however exactly the qualification is to be understood. But we badly need such a label if the difference I have in mind between psychoanalytic contributions to the Platonic–Aristotelian investigation and their 'virtue ethics' counterparts is to be made visible. So, when I say 'human excellences' I do *not* mean that as a synonym for 'human moral excellences', but rather as a more inclusive label for human excellences, understood so as to include both (for example) justice and fidelity to promises (assuming (p. 619) that's where the word 'moral' takes us, if it takes us anywhere) *and* (for example) resilience and proper ambition.

With our terminology thus fixed, we can observe that for 'virtue ethics', the excellences that matter are the core of justice, fidelity to promises, and so on. Psychoanalysis, on the other hand, has often explicitly resisted the idea that it's concerned with these or indeed with any moral virtues, with the possible exceptions of courage and truthfulness. In their place, it's common to find an emphasis in psychoanalysis on excellences (real or apparent) which, though seemingly excellences of character, bear a problematic relation to moral virtue, such as integration, or authenticity, and (in addition, or alternatively) on what might be called 'relational excellences': the capacity for other-relatedness, or good object relations, or the capacity to love. Even if a number of psychoanalytic thinkers have been as interested as was Aristotle in the proper development of our distinctively human natures, what they think that consists in is ostensibly quite different. The following remark of Loewald's exemplifies both the overlap and the divergence: 'a human being's becoming a person'—that's the overlap—is a matter not, indeed, of developing moral virtues but 'of the development of our love-life' (1978: 32). Again, D. W. Winnicott thought psychoanalysis has 'limited moral import' (Blass 2001: 207), but he didn't think that meant it was merely about health: 'the absence of psychoneurotic illness may be health but it is not life', and what psychoanalysis should offer is 'life' (Winnicott 1967: 360, cited by Blass 2003: 931).

In attempting to map the relations between psychoanalysis and 'virtue ethics', then, we have arrived so far at a three-way distinction, between those (if any) who see psychoanalysis as about human excellence or the proper development of our distinctively human natures *and* who see human excellence as consisting in the moral virtues (whatever exactly the list of those includes); those who see psychoanalysis as about human excellence or the proper development of our distinctively human natures but who see human excellence as consisting in some other set of characteristics; and those who see psychoanalysis not as about human excellence at all, but as about something much more modest—psychic health.

The picture is further complicated, however, by the fact that the phrase 'psychic health' itself stands for something different in different psychoanalytic writers. Two comparisons may serve to explain the difficulty. First of all Plato and Aristotle again: these philosophers might have been puzzled at the idea that if psychoanalysis is about psychic health, it cannot *also* be about human excellence, even human excellence interpreted as moral virtue. For Plato said that 'virtue is the health and beauty and well-being of the soul, and vice is the disease and weakness and deformity of the soul' (1961: 444 d–e), and

Aristotle something similar (1984: 1166a30–1166b1, 1166b25–1166b29). That is, in their theories, psychic health is either identical with human excellence or, perhaps, with some excellent distinctively human condition which underlies and explains it. The second comparison is with the notion of physical health. That notion is hardly straightforward. But for all that physical health is in some sense a condition we are *meant* to be in, and for all that it seems to resist reduction to a value-free notion of statistical normality (Boorse 1977; Kingma 2007), it is surely—and not just because non-human creatures (p. 620) too are meant to be physically healthy—implausible as an occupant of the exalted role Plato and Aristotle assigned to the health of the soul. For some psychoanalytic writers, then, psychic health is more like health of the soul in Plato and Aristotle—the developmental ideal for our distinctively human natures—whereas for others, it is more like physical health, a good, but a subordinate one (and which may relate only obliquely to ideals of character). Unfortunately, psychoanalytic writers may sometimes *say* it is more like the one, and treat it as more like the other. What is more, which one it *is* more like cannot always be inferred from the set of characteristics in which psychic health is said to consist.

Be that as it may, the more psychoanalysis treats psychic health like health of the soul for Plato and Aristotle, the more it looks as if psychoanalysis and 'virtue ethics' constitute attempts on the same Platonic–Aristotelian set of questions. Indeed, where psychoanalysis treats psychic health like health of the soul for Plato and Aristotle *and* where its focus is on (real or apparent) excellences distinct from the moral virtues—from justice, fidelity to promises, and so on—there is the possibility not only that they are attempts on the same set of questions, but rival attempts. Where by contrast psychoanalysis treats psychic health as more like physical health, this possibility tends to fade from view.

In this chapter I am going to do little more than fill out this sketch of the various theoretical options in a complex field.

Hartmann, Plato, and 'Psychic Health'

It has often been difficult for psychoanalysis to pin down a notion of psychic health that is not bound up with human excellences of one kind or another. A psychoanalyst who ostensibly tried to do so was Heinz Hartmann. Writing in the 1960s, Hartmann argued that:

> analysis is primarily a therapeutic and not a moral instrument … [W]e try to relieve a patient of his neurotic suffering, even if there is no chance that in the process he will become a morally better person. (1960: 69–70, 87)

> In his therapeutic work [the analyst] will keep other values in abeyance and concentrate on the realization of one category of values only: health values. (1960: 55)

And Hartmann opposed what he called 'health ethics', which makes:

> an equation of moral badness with mental dysfunction … [hoping that] in attacking neurosis individual therapy can also do away with everything that is considered 'bad' in human nature. (1960: 70)

Hartmann is thus ostensibly defending a conception of psychic health which is independent of any human excellences (moral or otherwise), because we can state what it is (p. 621) to be psychically healthy without reference to 'social, moral and other values' (1960: 55), or indeed to any value beyond health itself.

Notice, however, that in the last passage quoted, Hartmann has 'bad' in scare-quotes—as if he might mean not what's really bad, but what society (in its drearily conventional way) merely *calls* 'bad'. If that is so, then to be sure we could be free of neurosis (and so presumably at least minimally psychically healthy) while remaining 'bad', that is, while straying from conventional moral norms. Such a line of thought is also present in Freud: some neurotic illness, Freud said, could be prevented if society took thought as to what was really worth proscribing in the way of sexual behaviour, the clear implication being that in this area at least, much of what we call 'wrong' isn't really wrong (1962: 278; cp. 1963b: 434). 'Psychoneuroses spring from the sexual needs of people who are unsatisfied', Freud said (1959: 186), while at the same time favourably contrasting 'strong natures who openly oppose the demands of civilization' with 'well-behaved weaklings' (1959: 192). However, the fact that psychic health is compatible with (scare-quoted) 'badness'—that is, with the absence of merely conventional excellences like sexual abstinence outside marriage—doesn't show it is compatible with (real not scare-quoted) badness. But if psychic health is *not* compatible with real badness, that is as much as to say that it implies real rather than merely conventional human excellence, whatever exactly that turns out to be.

Moreover, notwithstanding Hartmann's claim that psychoanalysis aims simply at freedom from neurosis, it also in his view promotes 'more successful integration' and improved 'communication between the superego and the ego' (1960: 86); though it does not seek to promote particular moral values, it is interested in these values 'in regard to their authenticity in expressing the complex and more or less integrated dynamics of … moral personality' (1960: 51). So explained, integration and authenticity are conceptions of the proper relations between different parts of the psyche. Anyone familiar with psychoanalysis will recognize these as extremely common ways in which psychoanalysis glosses the notion of psychic health. And remember Aristotle's claim that the virtuous man's 'opinions are harmonious, and he desires the same things with all his soul' (1984: 1166a30–1166b1). If integration is not actually one of the *virtues* in Plato's and Aristotle's catalogues, it is surely the good state of the soul which may be said to underlie virtue. So we should beware of taking Hartmann's claim to have described a notion of psychic health that's independent of 'social, moral and other values' too much at face value.

Psychic Health, Character Disorders, and Neurosis

At this point, perhaps unexpectedly, the history of psychoanalysis comes to play a part in the investigation. Since at least the 1950s, it has often been remarked that analysands tend to present not with hysteria or obsessional neurosis—seemingly the staples of (p. 622) Freud's own practice, and certainly the stuff of many of his most famous case histories—but something more diffuse: 'narcissism', or an inability to achieve depth in relationships, or 'a general sense of not being able to get along' (Giovacchini 1975: 31).[4] As Allen Wheelis put it in 1958:

> [a] change in the patterns of neuroses [fell] within the personal experience of older psychoanalysts, [while younger ones] became aware of it from the discrepancy between the older descriptions of neuroses and the problems presented by the patients who come daily to their offices. *The change is from symptom neuroses to character disorders.* (Lasch 1980: 42, my italics)

Why this shift occurred is not clear. Perhaps large-scale cultural changes meant that society simply ceased to produce neurotics, and produced something else instead (Lasch 1980). Perhaps on the other hand what changed is not the epidemiology of these conditions themselves but the client base of psychoanalysis, thanks for example to its increasing cultural acceptance, or to the extent to which analysts were employed outside psychiatric institutions (Horwitz 2002).

Whatever the truth of the matter, the mere fact of the historical shift bears on the account we are to give of psychic health as the notion is deployed in psychoanalysis. An attractive (though not uncontroversial) way of explaining the notion of psychic health on analogy with that of physical health is in terms of the absence of symptoms. Since Freud defined the neuroses in terms of their symptoms, absence of symptoms at least makes sense as a formulation of psychic health, as early psychoanalysis took itself to be pursuing it. But it is useless as a conception against which to interrogate psychoanalysis since the 1950s—or whenever exactly the relevant shift took place. For it is built into the characterization of post-1950s analysands that what they brought to analysis were *not* symptoms. Presumably they came to analysis for a reason, though, and yet, relative to the 'symptom-freedom' definition of psychic health, when they began analysis they were already perfectly well. What's more, if we find mid-century psychoanalytic writers such as Hartmann *claiming* that psychic health is merely freedom from 'neurotic disease', and yet —when one scratches the surface—tacitly thinking of psychic health in terms of human excellence, or of the proper arrangement of parts of the psyche—one explanation for this is that psychoanalysis in Hartmann's time had not caught up with itself, and was affixing labels that fitted the pre-1950s scene to theories suited to the new generation of analysands.

However, the very fact that it is historical reasons that underlie the possibility of thus reading Hartmann's notion of psychic health against itself suggests it would be a mistake (p. 623) to conclude that *wherever* psychic health, characterized negatively as symptom-freedom, appears to be the animating ideal of psychoanalysis what is really in play is some characterological notion. If the notion of psychic health deployed in the neurotics' heyday is no more *merely* a health conception—no more free of 'social and moral' notions—than Hartmann's, that needs to be separately argued for. The obvious place to look in order to put that suggestion to the test is, of course, the work of Freud himself, as I do in the next section.

Neurosis, Psychosis, and Freud's 'Work Ethic'

The following passage, from 'The Loss of Reality in Neurosis and Psychosis' (Freud 1961: 183–7), can stand as well as any as a summary of Freud's view of psychic health as poised between neurosis and psychosis:

> Both neurosis and psychosis are … the expression of a rebellion on the part of the id against the external world, of its unwillingness … to adapt itself to the exigencies of reality … In neurosis a piece of reality is avoided by a sort of flight, whereas in psychosis it is remodelled …. We call behaviour 'normal' or 'healthy' if it combines certain features of both, … [i.e.] if it disavows reality as little as does a neurosis, but if it then exerts itself, as does a psychosis, to effect an alteration in it…. [T]his expedient, normal behaviour leads to work being carried out on the external world; it does not stop, as in psychosis, at effecting internal changes…. [It is] *alloplastic*.

Neurosis and psychosis, that is, have—according to Freud—something in common, in that each involves a denial of reality (in the case of neurosis, of an aspect of inner reality, in psychosis of an aspect of the external world). So in the second we find delusion, and in the first, symptom-formation. Psychic health, meanwhile, has something in common with both. It has something in common with neurosis in that *external* reality is not denied in either case. It has something in common with psychosis in that in each case the subject goes to work on and transforms something—in psychosis, the subject's inner world; in health, external reality itself. The thought is echoed in many other remarks, spread widely across the corpus of Freud's writings.

Two features of this conception deserve comment here. First, what underpins psychic health—that is, what explains *why* a given person turns neither to neurosis nor to psychosis, when they don't—is, though Freud doesn't use this word, a human excellence: roughly, the excellence of being able to look reality in the eye, disagreeable though it may be. In the passage quoted above, admittedly, it is *the id's* unwillingness to adapt itself to the exigencies of reality, but in other related passages Freud uses more familiar moral (p. 624) language: 'Today neurosis takes the place of monasteries which used to be the refuge of all whom life had disappointed or *who felt too weak to face it*' (1957: 50, my italics); neurotics 'turn away from reality because *they* [not just their ids] find it

unbearable' (1958: 218). Nor should it be a surprise to find Freud holding such a view. 'The great ethical element in psychoanalytic work', as he acknowledged, 'is truth and again truth … Courage and truth are of what [sic] … [people] are mostly deficient' (Hale 1971: 170). Courage and truthfulness—which I take it is what he meant, writing in English, by 'truth'—are unarguably human excellences (and indeed excellences in Aristotle's catalogue), and these seem to be just the characteristics people 'too weak to face reality' don't have.[6, 7]

Beyond this, Freud's characterization of psychic health draws heavily upon a second characteristic, namely the determination to work or to effect change in the world (to be 'alloplastic'). Freud sees the 'healthy' person as one who seeks to *shape* reality, giving form to his (and it always is *his*) wishes. (Here Freud is surely influenced by Nietzsche: see Nietzsche 1882/2001: 163–4; Ridley 1998: 136.) 'The energetic and successful man', Freud says, 'is one who succeeds by his efforts in turning his wishful fantasies into reality' (1957: 50). 'Scientific work' enables one to 'gain some knowledge about the reality of the world, by means of which we can increase our power and in accordance with which we can arrange our life' (1964a: 55). 'The man of action will never give up the external world on which he can try his strength' (1964b: 84). By joining with others, and 'guided by science', we can 'go over to the attack against nature and subject her to the human will' (1964b: 72).[8]

What we might call Freud's 'work ethic' is also to the fore in some of his remarks on sexuality. He claims that the child 'is meant to grow up into a strong and capable person with vigorous sexual needs and to accomplish during his life all the things that human beings are urged to do by their instincts' (1953: 223). One might think, on Freudian principles, that what one is 'urged to do by instinct' is simply to satisfy the instinct (1953: 168). However, Freud condemns masturbation—arguably the equal of non-masturbatory sexual activity as far as satisfying instinct is concerned—on the grounds that 'it teaches people to achieve important aims without taking trouble and by easy paths instead of through an energetic exertion of force' (1959: 199). Sexual intercourse is better than (p. 625) masturbation, in other words, not because it is better at satisfying instinct but because it is a form of 'work'.

Though Freud connects the value of work with the virtue of being able to look reality in the eye—as when he says of 'the ordinary and professional work that [because it requires no extraordinary talents] is open to everyone' that 'no other technique for the conduct of life attaches the individual so firmly to reality' (1964b: 80 n.1)—the value Freud places on 'work' is clearly additional to the value he places on facing reality, or on courage and truthfulness. For one could face reality simply by passively suffering whatever reality throws in one's way, paying without murmur the price in frustration. 'Going over to the attack against nature', on the other hand, is something extra. Moreover, insofar as neurosis, in Freud's view, is a 'flight from reality', 'work' is more than is required for psychic health as long as that means freedom from neurosis, since the courageous passive sufferer (or indeed the seeker of autoerotic sexual satisfaction) would no more be flying from reality than would the person who 'goes over to the attack'. It would seem,

then, that in Freud we have a minimal conception of psychic health (as freedom from symptoms), and a further conception which, though Freud also *calls* it 'health'—'remaining in lasting possession of health and efficiency' (1959: 181), for example—is a more substantive ideal of 'living life to the full' (compare Winnicott's 'life'). But both are bound up with notions of human excellence: the more substantive ideal obviously so, because it relies on Freud's 'work ethic', but the minimal conception too, since even it requires the excellences of courage and truthfulness.

Object Relations and the 'Relational Excellences'

Notwithstanding their sometimes medical self-presentation, both Hartmann's and Freud's conceptions of psychic health therefore build in human excellences, albeit in different ways. Something very similar is true of a final leading strand in psychoanalysis I will now consider.

A great many versions of post-Freudian psychoanalysis are organized by a master distinction—the distinction between 'immature' and 'mature dependence' in Winnicott, for example, or 'mergedness' and 'separateness' (Mahler 1986; Winnicott 1985, 1990), 'undifferentiation' and 'differentiation' (Loewald 1978), 'fusion' and 'autonomy' (Mitchell 1988), and the 'paranoid–schizoid' and 'depressive' positions (Klein 1988). Notwithstanding the variations in terminology, all these distinctions are designed to function as the measure of psychic health or, alternatively, of psychopathology. (They are also designed to function as the measure of psychological maturity, but that is another story: see Harcourt 2018a.) As Margaret Mahler puts it (1986: 222), 'the object-relationship'—meaning the degree of 'separation-individuation'—'is the most reliable single factor by which we are able to determine the level of mental health'. Again, the central 'pathology' (p. 626) which so much post-Freudian psychoanalysis sets out to address—'narcissism'—is frequently explained as a way of relating to others that is 'pathological' precisely insofar as it is 'merged' or 'undifferentiated' or 'dependent' (Eagle 1984; Fonagy 1999: 602; Lear 1998: 134; Loewald 1978).

Undifferentiation, 'narcissism' and so on are often described as 'boundary issues': according to Kohut, for example, the 'intimacy of [pathologically] immature … relationships' involves the 'blurring of boundaries between selves' (Chazan 1998: 74). But what does talk of boundaries between people really amount to? Plausibly, blurring of boundaries refers to that state of affairs when we become over-identified with another person so that we cannot think accurately about what they think or want: when we ask ourselves about them, what we come up with is just an echo of our own thoughts and wants. Or, alternatively, it refers to that state of affairs when we can't think accurately about ourselves: we have been so taken over by another person that when we try to consult our own thoughts or wants, all we come up with is what they think or want. The terminology of undifferentiated/differentiated (and so on), that is, marks a contrast between on the one hand something like egotism, incapacity for other-concern, failure to

be alive to others as independent centres of thought, feeling, and initiative—'narcissism' being at least an example of and perhaps a catch-all term for all of these defects; and (on the other), in Betty Joseph's words, 'real relating, in which the other'—and oneself—are 'thought of as separate' (2004: 161). But surely the capacity for 'real relating, in which the other is thought of as separate' is a human excellence, whether or not it is a moral one (Harcourt 2018a). (It might even be an Aristotelian virtue—*philia*, usually rendered 'friendship'—even if that virtue is not conceived as Aristotle conceived it.)

Psychoanalysis and Ideals of Character

The literature of psychoanalysis is vast and I cannot aspire in one chapter to cover all of it. But at least three readily recognizable psychoanalytic conceptualizations of psychic health—which may of course also overlap, in ways I have not explored—turn out, sometimes reading the psychoanalysts against themselves, sometimes simply taking them at their word, to involve ideas of human excellence, or excellence of character, whether or not these are also moral: courage, truthfulness, and 'work' in Freud; authenticity and, especially, integration in Hartmann (and many others); and the relational excellences summarized by 'real relating in which the other is thought of as separate' in Joseph herself, and in Mahler, Winnicott, and a great many others.

This might not seem like a very exciting result. Relative to the aspiration of some analysts to be 'doing medicine', and therefore to be doing something 'value-free'—if indeed somatic medicine *is* that—of course the result is worthwhile, but that aspiration seems to be driven, as much as anything, by an institutional need from within psychoanalysis (p. 627) to seem professionally respectable plus a typical twentieth-century anxiety that being mixed up with 'values' is *not* respectable. Certainly the thought that *mental* health—of which the psychoanalysts' 'psychic' health is presumably an aspect—is inextricably tied to evaluative notions is familiar, even if not universally accepted (Horwitz 2002; Horwitz and Wakefield 2007).

A more interesting question is what this involvement of characterological notions in psychoanalytic conceptions of psychic health really means.

Here the history of psychoanalysis is relevant once again. If the typical post-1950s psychoanalytic presentation is not with symptoms but with a 'character disorder', then it shouldn't come as a surprise that—typically—psychoanalysts' accounts of what psychoanalysis is about should have moved away from absence of symptoms towards something more like a good state of character. Thus R. D. Hinshelwood writes that though psychoanalysis initially aimed 'simply at the relief of symptoms through interpreting dream symbols', it then 'became the analysis of character, [and so] *changed from a mere medical technique to a moral practice*' (1997: 3, my italics). However, we should be as careful about taking at face value Hinshelwood's claim that psychoanalysis is a 'moral practice' as we should about taking at face value Freud's that it is a 'medical procedure', or Hartmann's that it 'concentrate[s] on the realization of one category of values only: health values'. For the question is surely not only whether the psychoanalytic conception

of psychic health is about freedom from symptoms or rather about some condition of character, be it 'integration' or courage and truthfulness or 'good object relations': if that *were* the only question, it would be very easy to answer.

I suggest that the interesting question in this area breaks down into two. One is whether, insofar as psychoanalytic conceptions of psychic health build in characterological notions, psychic health plays the ambitious role in psychoanalysis that 'health of the soul' plays in Plato and Aristotle (that is, as a conception of the proper development of our distinctively human natures), or some more modest role (compare again physical health). The other question, which is orthogonal to the first, concerns the relationship between the characterological notions implicated in the conceptions of psychic health under review, and the notions of virtue or human excellence we find in Plato and Aristotle and, more especially perhaps, among their contemporary followers in 'virtue ethics'. To see why this question *is* orthogonal to the first, remember that in Plato and Aristotle, psychic health is both an ambitious notion—the end of a distinctively human developmental trajectory— *and* consists in a set of virtues containing at least the core of fidelity to promises, justice, and so on. There is room therefore for psychoanalytic conceptions of psychic health, for all that they seem inescapably to build in characterological notions, to relate to this Platonic–Aristotelian template in more than one way. In particular, psychic health as conceived by psychoanalysis could thus be the end of a distinctively human developmental trajectory *and* be realized in moral virtue, i.e. human excellence understood so as to include the core of fidelity to promises etc. itself: on this view, psychoanalysis would track Plato and Aristotle quite closely, and some contemporary 'virtue ethics' still more so. Or it could be the end of a distinctively human developmental trajectory but consist in excellences distinct from those moral ones (though perhaps overlapping them). Or again, psychic health might not be the end of (p. 628) a distinctively human developmental trajectory at all, but represent, like physical health, a more modest good.

In the two final sections, I make a start on a reasoned comparison between these three views.

Psychic Health, Proper Human Development, and the Virtues

The first of the above views—psychic health as the end of a distinctively human developmental trajectory *and* realized in moral virtue—features in a worked-out form at least in psychoanalysts Erich Fromm (1947), Karen Horney (1950/1991), and Erik Erikson (1964); cp. also Narvaez, this volume. Citing Aristotle, Fromm says not only that 'in order to know what is good for man we have to know man's nature' but also that 'virtue is … identical with the realization of man's nature' (1947/2003: 12, 19). He then appeals to the notion of human nature's proper or, alternatively, improper development to explain the difference between the real excellences of being 'accepting', 'devoted', 'self-confident', or 'economical', and the superficially similar but non-excellent traits of being 'passive',

'submissive', 'arrogant', or 'stingy' (1947/2003: 82). Moreover the difference between them is not just that between real excellences and counterfeits, but that the real ones flow, as the impostors do not, from a 'total character organization'—'productiveness' (1947/2003: 42, 60)—in which psychic health consists. This 'organization' is modelled on—and here the Aristotelian teleology meets the Freudian—Freud's (early) conception of proper development as the achievement of 'full genitality', though Fromm's notion, he says, is not one of merely sexual (re)production, but of a total 'mode of relatedness' to the world (1947/2003: 61).

Unsurprisingly, then, a contemporary neo-Aristotelian might be drawn to this first view, the neatest possible coincidence between psychoanalysis (in some versions) and 'virtue ethics'.[9] On that view, would psychoanalysis have anything to add to 'virtue ethics', beyond endorsement from afar? Potentially yes. Plato's and Aristotle's accounts of virtue and vice look as if they have two levels, the level of the virtues and vices themselves and an underlying moral psychology, with the underlying moral psychology—the interrelation, healthy or otherwise, of the different parts of the soul—supposedly explaining the presence or absence of virtue (Irwin 2007: 154). A plausible complaint against at least Plato, however (Lear 2003; Williams 2000), is that what it is for the parts of the soul to be healthily related is fixed in advance by the dominance of reason in his conception of virtue, so the claim that virtue is the health of the soul is no more than an (p. 629) honorific relabelling. Whether this complaint is fair or not isn't my business here—and I do not have the space to ask whether (for example) Fromm's 'productiveness' itself can do better. The point is that if there is something for psychoanalysis to add to 'virtue ethics' just here, it is surely this: to give a genuinely independent characterization of the underlying level which *leaves conceptual room* for nasty surprises (so the concepts of psychic health and of moral excellence, conceived so as to include the core of justice and so on, are distinct), but at the same time shows there aren't any.[10]

Is this first view over-optimistic? If there can be people who are psychically healthy—relative to the conception of psychic health under discussion—who do not also possess the virtues of justice, fidelity to promises and so on, then the first view is over-optimistic since it precisely denies that possibility. So assessing whether it *is* over-optimistic breaks down into as many distinct tasks as there are conceptions of psychic health (I reviewed three).

Let's begin with Freud. If we take seriously the teleological language of claims such as '[the child] *is meant to* grow up into a strong and capable person . . .' (1953: 223, italics mine), Freud seems to be saying of being 'strong and capable' (etc.) what Plato and Aristotle say of possession of the virtues, including justice and so on, namely that it constitutes the proper development of our natures, given the kinds of creatures we are. But the relationship between being strong and capable (etc.) and the moral virtues is—*pace* Fromm, who of course borrowed his ideal of 'productiveness' from this bit of Freud—not straightforward. Since 'work' is arguably a virtue, and if we assume that Freud's *minimal* psychic health is a condition for 'being strong and capable' (etc.) and minimal

psychic health implies courage and truthfulness, then 'being strong and capable' (etc.) guarantees three virtues. The tricky question is whether it guarantees all the others.

It may look as if Freud thought it guaranteed many others: 'Anyone who has succeeded in educating himself to the truth about himself', he said, 'is permanently defended against the danger of immorality' (1963b: 434). However, the context suggests that what he means is that anyone who chooses to violate merely conventional norms of sexual behaviour in pursuit of instinctual satisfaction won't count as immoral as far as *that* is concerned (whether 'after the completion of their treatment' they opt for 'living a full life', 'absolute asceticism', or something in between). He just doesn't seem to have been thinking about justice or fidelity to promises. But now couldn't someone strong and capable—a politician or (favourite example of Freud's) an artist—be deficient in these virtues? Surely the human record suggests that they could.

Now let's turn to psychic health as 'real relating', or 'good object relations'. Some psychoanalytic writers—and of course not only they—have been extremely optimistic about what 'real relating' can accomplish: to assume that if we get it right in our 'love lives' (in Loewald's sense) all other forms of human goodness will flow from that (Sagan 1988: 194). But it does not seem probable that love is identical with the dispositions in which justice, fidelity to promises, or kindness—to name a few—consist. For at least (p. 630) these virtues are such that we owe them indifferently to everybody, whereas the relational excellences are such that we owe their exercise only to those to whom we already stand in a particular kind of relationship. So these virtues seem to have different psychological sources, which would explain why a person can plausibly be very good in respect of the one and highly indifferent in respect of the others (such as the exemplary good neighbour who steers clear of intimate relationships 'because he is not very good at them'). Indeed Freud himself said as much: love does well at explaining good behaviour between people who stand in special relations to one another. But if 'real relating' is to underwrite justice, fidelity to promises and so on, we need it to do the same explanatory work for all persons. But we don't love everybody, so why should it? (Eros 'binds in love among a loving pair', but 'civilization depends on relations between a considerable number of individuals', 1964b: 108.) At best, if 'real relating, in which the other is thought of as separate' *is* a kind of master-virtue which implies all the others, that case has yet to be made out.

Finally, how about integration? Here social context is important since integration, though possibly not a wholly 'individualistic' notion, seems less context-dependent than some of the core moral virtues. As Roger Money-Kyrle (1952) argues, psychoanalysis may be able to show that there is a functional difference between the 'authoritarian' and the 'humanistic' conscience, and that one is psychically healthier than the other. (The inferiority of the authoritarian conscience in respect of psychic health consists in the fact that in that formation ego and superego are less well integrated with each other.) But, importantly, he adds that 'the humanistic consciences of different individuals are by no means necessarily identical. They have a common form or pattern which distinguishes them sharply from the authoritarian type ... but the content ... may be widely

different' (1952: 233). One reason for this may simply be differences in individual judgement: not everybody with (say) a humanistic conscience will have identical moral attitudes, so—on reasonable assumptions about the possibility of moral error—at least some of these moral attitudes will be mistaken. But from the perspective of psychic health, the correct ones and the mistaken ones are on all fours, so there can be moral differences which differences in psychic health do not explain. Another reason is the difference in cultural resources available to different individuals. One cause of the fact that two well-integrated individuals may differ with respect to their moral attitudes lies in the difference in resources which their respective social contexts make available to them —if one lives in an apartheid system and the other in a liberal democracy, for example. It would be consoling to think that everyone who lives under the former system must be psychically unhealthy, but even if this is true of the system's architects it is very unlikely to be true of everyone favoured by the system and living comfortably under it, and all the more so the more the system beds down and comes (to them) to seem normal. And yet some of the normal attitudes under the system are bad, so the person with the 'humanistic' conscience under the apartheid system and his or her counterpart under the liberal democracy are surely not indistinguishable in point of moral virtue: under the former system at least some virtues—most obviously justice—are **(p. 631)** impossible for the privileged conformist. But to say that such a person cannot be psychically healthy because they are not virtuous is question-begging; to say they must be virtuous because they are psychically healthy is manifestly wrong.

In sum, it looks as if on all three psychoanalytic conceptions of psychic health we have considered, it is possible to be psychically healthy but not to possess the virtues of justice, fidelity to promises, and so on. So the view that psychic health is the end of a distinctively human developmental trajectory *and* is realized in moral virtue is indeed over-optimistic.

Psychic Health With and Without Teleology

If that view is over-optimistic, to what extent do the conceptions of psychic health under review constitute a challenge to those conceptions of the proper development of our natures which tend to dominate 'virtue ethics', and which emphasize the centrality to proper human development of justice, fidelity to promises, and other core moral virtues? This depends on at least two further questions: first, as already noted, it depends on whether psychic health is itself seen as the end of a distinctively human developmental trajectory or in a more modest light. And secondly, if psychic health is seen ambitiously rather than modestly, it depends further on whether psychic health (and the non-moral presumptive excellences it involves) is seen as *the only* end of proper human development —call this the exclusive view—or as *one* such end (the inclusive view). I deal with the second of these issues first.

The exclusive view comes at a cost. Philosophers who think there is such a thing as the proper development of our natures will presumably also hold that a test of whether a

characteristic really *is* a human excellence is whether it features on the list of characteristics a human being would possess if their natures were properly developed, however exactly that notion is to be explained in its turn. So an adherent of the exclusive view, believing that the core moral virtues are not what the proper development of our natures consists in, faces a choice. Either they have to hold that some or all of these virtues are indeed good characteristics, but explain their goodness otherwise than by appeal to the notion of proper development—surely an uncomfortable position given that this is how they explain the goodness of psychic health; or else, if that's the only way goodness is to be explained, they must conclude that the core moral virtues aren't really good characteristics, but only conventionally regarded as such. On the latter alternative, the exclusive view potentially underwrites a scepticism about the value of the moral virtues—compare of course Nietzsche—which will be familiar to readers of (some) psychoanalysis. That may be a successful line of critique against, for example, an old-fashioned conventional sexual morality. But I take it that it is not an attractive option to treat justice, or (p. 632) fidelity to promises, or generosity, as making unreal demands on us, or as merely conventional virtues—even if we may sometimes behave in accordance with them for the wrong reasons (such as timidity or conformism). Lacan seems to have thought otherwise ('a form of ethical judgment is possible which gives this question the force of a Last Judgment: Have you acted in conformity with the desire that is in you?', 1992: 314). But I suspect that the number of psychoanalysts clear-sightedly to have embraced the exclusive view are few. Those psychoanalysts who have seen psychic health as the end of a distinctively human developmental trajectory, but with an at least ambiguous relation to the moral virtues—Winnicott, or Loewald (1978: 5–6); Anna Freud (1998) with her notion of 'developmental lines' would be another example—have rather tended either to overestimate what flows from personal love (cp. Backström, this volume), or not to have been very interested (at any rate professionally) in virtues such as justice and fidelity to promises, and so to have left the potentially anti-moral implications of their alternative non-moral teleology unexplored. (I would also say that a great many of the popular moralists of the twentieth and twenty-first centuries, where a 'therapeutic teleology' has been almost a default, are, for one or other of these reasons, in the same unwitting predicament.) Since psychoanalysts are presumably no less eager not to be let down, lied to, exploited, or insulted than the next person, it is a question well worth further investigation why the advocates of 'life' (Winnicott 1967: 360) should sometimes have been so dismissive of 'morals' (Winnicott 1990b).

However that may be, the notion of 'proper development' doesn't require that we think of the core moral virtues or, alternatively, a set of non-moral excellences as exclusive claimants to the title of the proper development of our natures: human beings may by nature be such that more than one set of excellences constitute *a* proper development of their natures. So, supposing that psychic health didn't imply enough moral virtues, it would not follow that the moral virtues aren't aspects of the proper development of our natures, or were merely apparently good: it might rather be that not all genuine excellences can be instantiated in the life of a single albeit properly developed human being (Williams 1985: 153). Still, given the possibilities for value-conflict on this picture,

and the emphasis on freedom from conflict both in (neo-)Aristotelian ethics (in the shape of the thesis that any virtue implies every virtue (Cooper 1998)) and in many strains within psychoanalysis,[11] it is perhaps easy to see why those interested in reconciling psychoanalysis with those conceptions of the proper development of our natures which tend to dominate 'virtue ethics' may have tended to prefer the over-optimistic view I have ascribed to Fromm and Horney.

On an alternative conception, psychic health would not be identified with the proper development of our natures at all, whether that consists in the moral virtues, in some alternative set of excellences, or in both. The goals of psychoanalysis, on this view, would thus be more modest, notwithstanding the fact that psychic health builds in (p. 633) characterological components on several different psychoanalytic conceptions of it. Still, because nothing in psychic health (as modestly conceived) is inconsistent with virtue, psychic health would still be pictured not indeed *as* the proper development of our natures, or as a sufficient condition of that (as on the over-optimistic view), but as a necessary condition. As Freud says in a letter to J. J. Putnam, 'I quite agree … that psychoanalytic treatment should find its place among the methods whose aim is to bring about the highest ethical and intellectual development of the individual. Our difference … is confined to the fact that I do not wish to entrust this further development to the psychoanalyst' (Hale 1971: 170–1). Or as R. D. Hinshelwood puts it in a different idiom, psychoanalysis aims at 'greater integration' because that is a precondition of autonomy, rationality, and 'holding together intrapsychically … various choices'. But though integration is a 'route on which a person must travel to become a moral being', it is not identical to becoming one (1997: 154–5; cp. Hursthouse 1999: 207). Psychoanalysis would thus play a part in a person's moral education, but because how much further we take that is up to us, psychoanalysis might complete its work while still leaving 'a person of but moderate worth' (Freud 1961: 119).

But is psychic health even a necessary condition for moral virtue, or indeed for a range of non-moral human excellences? If the absence of psychic health can accompany moral virtue, the answer is no. Were the topic *mental* health as understood in psychiatry, rather than psychic health as understood in psychoanalysis, it would surely be clear that its absence *can* accompany moral virtue: the severely mentally ill can be as kind, honest, or fair as the next person even if their illnesses sometimes prevent these dispositions from being actualized. To stick with psychic health in psychoanalysis, however, the absence of psychic health can of course accompany moral virtue if its absence is itself *needed* for moral virtue. But now, might not caring for or healing others *for the right reasons* be an imaginative way of dealing with one's own need to be cared for or healed? Because it is done for the right reasons the caring really is virtue, not some behaviourally similar but otherwise motivated counterfeit (contrast Fromm 1947/2003); but because it's an imaginative negotiation of a psychic problem, in *that* virtuous person at least, virtue requires the presence of the problem. The question is whether what I am blandly calling 'the problem' is an instance of psychic ill-health. Psychoanalysis—for better or worse— often points to the same mechanisms as underlying the normal and the pathological, as in Klein's deployment of the notion of projective identification, or Anna Freud's of

projection, all the work of distinguishing between normal and pathological being left to terms like 'normal' and 'excessive' (e.g. A. Freud 1998: 25; Klein 1988). Still, if the very presence of knots in the psyche *is* a mark of psychic ill-health—as some psychoanalytic talk of the 'fully analysed personality' might lead one to imagine—then psychic health is not even necessary for virtue, since these knots belong to the material out of which the human psyche manufactures virtue (or for that matter humour, or art). On the other hand, perhaps the fact that it's *virtue* rather than (say) symptoms that is manufactured shows that there is no pathology. So the claim that virtue is possible without psychic health, modestly conceived, has yet to be made out.

References

Aristotle (1984). *Nicomachean Ethics*. In J. Barnes (ed.), *Complete Works of Aristotle*, Vol. 2: The Revised Oxford Translation (Bollingen online edition) (pp. 1729–1867). Princeton, NJ: Princeton University Press.

Bateman, A. and Holmes, J. (1995). *Introduction to Psychoanalysis: Contemporary Theory and Practice*. London: Routledge.

Blass, R. (2001). 'On the ethical and evaluative nature of developmental models in psychoanalysis'. *The Psychoanalytic Study of the Child 56*: 193–218.

Blass, R. (2003). 'On ethical issues at the foundation of the debate over the goals of psychoanalysis'. *The International Journal of Psychoanalysis 84*: 929–943.

Boorse, C. (1977). 'Health as a theoretical concept'. *Journal of the Philosophy of Science 44*: 542–573.

Chazan, P. (1998). *The Moral Self*. London: Routledge.

Cooper, J. M. (1998). 'The unity of virtue'. *Social Philosophy and Policy 15*: 233–274.

Eagle, M. (1984). *Recent Developments in Psychoanalysis*. Cambridge, MA: Harvard University Press.

Erikson, E. (1964). *Insight and Responsibility*. London: Faber.

Fonagy, P. (1999). 'Psychoanalytic theory from the point of view of attachment theory and research'. In J. Cassidy and P. Shaver (eds), *Handbook of Attachment*. New York/London: Guilford Press.

Fonagy, P. (2003). 'The Klein-Bion model'. In P. Fonagy and M. Target, *Psychoanalytic Theories: Perspectives from Developmental Psychology*. London: Whurr.

Freud, A. (1998). 'A psychoanalytic view of developmental psychopathology'. In R. Ekins and R. Freeman (eds), *Anna Freud: Selected Writings*. Harmondsworth: Penguin.

Freud, S. (1953). 'Three essays on the theory of sexuality' [1905]. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 7 (pp. 125–243). London: Hogarth Press/Institute of Psychoanalysis.

Freud, S. (1955). 'Studies on hysteria' [1893–95]. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 2. London: Hogarth Press/Institute of Psychoanalysis.

Freud, S. (1957). 'Five lectures on psychoanalysis' [1910]. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 11 (pp. 3–55). London: Hogarth Press/Institute of Psychoanalysis.

Freud, S. (1958). 'Formulations on the two principles of mental functioning' [1911]. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 12 (pp. 213–226). London: Hogarth Press/Institute of Psychoanalysis.

Freud, S. (1959). ' "Civilized" sexual morality and modern nervous illness' [1908]. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 9 (pp. 177–204). London: Hogarth Press/Institute of Psychoanalysis.

Freud, S. (1960). 'New introductory lectures on psycho-analysis' [1933]. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 22 (pp. 3–158). London: Hogarth Press/Institute of Psychoanalysis.

Freud, S. (1961). 'The loss of reality in neurosis and psychosis' [1924]. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 19 (pp. 183–190). London: Hogarth Press/Institute of Psychoanalysis. (p. 635)

Freud, S. (1962). 'Sexuality in the aetiology of the neuroses' [1898]. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 3 (pp. 261–286). London: Hogarth Press/Institute of Psychoanalysis.

Freud, S. (1963a). 'Introductory lectures on psychoanalysis' [1916–17]. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 15. London: Hogarth Press/Institute of Psychoanalysis.

Freud, S. (1963b). 'Introductory lectures on psychoanalysis' [1916–17]. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 16. London: Hogarth Press/Institute of Psychoanalysis.

Freud, S. (1964a). 'The future of an illusion' [1927]. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 21 (pp. 5–58). London: Hogarth Press/Institute of Psychoanalysis.

Freud, S. (1964b). 'Civilization and its discontents' [1930]. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol 21 (pp. 64–148). London: Hogarth Press/Institute of Psychoanalysis.

Fromm, E. (1947). *Man for Himself*. Greenwich, CT: Fawcett.

Giovacchini, P. (1975). *Psychoanalysis of Character Disorders*. New York: Jason Aronson.

Hale, N. G., Jr. (1971). *James Jackson Putnam and Psychoanalysis: Letters between Putnam and Sigmund Freud and Others*. Cambridge, MA: Harvard University Press.

Harcourt, E. (2015). 'The place of psychoanalysis in the history of ethics'. *Journal of Moral Philosophy 12*: 598–618.

Harcourt, E. (2018a). 'Madness, badness, and immaturity: some conceptual issues in psychoanalysis and psychotherapy'. *Philosophy, Psychiatry and Psychology 25*(2): 123–136.

Harcourt, E. (2018b). 'Psychoanalysis, the good life, and human development'. *Philosophy, Psychiatry and Psychology 25*(2): 143–147.

Hartmann, H. (1960). *Psychoanalysis and Moral Values*. New York: International Universities Press.

Hinshelwood, R. D. (1997). *Therapy or Coercion?* London: Karnac.

Horney, K. (1950/1991). *Neurosis and Human Growth*. New York: Norton.

Horwitz, A. (2002). *Creating Mental Illness*. Chicago, IL: University of Chicago Press.

Horwitz, A. and Wakefield, J. (2007). *The Loss of Sadness*. Oxford: Oxford University Press.

Hursthouse, R. (1999). *On Virtue Ethics*. Oxford: Oxford University Press.

Hursthouse, R. and Pettigrove, G. (2016). 'Virtue ethics'. In *The Stanford Encyclopedia of Philosophy*, ed. E. Zalta, **https://plato.stanford.edu/entries/ethics-virtue/**.

Irwin, T. H. (2007). *The Development of Ethics*. Vol. 1: *From Socrates to the Reformation*. Oxford: Oxford University Press.

Joseph, B. (2004). ' "Where there is no vision": from sexualization to sexuality'. In D. Bell (ed.), *Reason and Passion: A Celebration of the Work of Hanna Segal*. London/New York: Karnac.

Kingma, E. (2007). 'What is it to be healthy?' *Analysis 67*: 128–133.

Klein, M. (1988). 'Explanatory notes'. In *Envy and Gratitude*. London: Virago.

Lacan, J. (1992). *The Seminar of Jacques Lacan VII: The Ethics of Psychoanalysis*, trans. D. Porter. New York: Norton.

Lasch, C. (1980). *The Culture of Narcissism*. London: Sphere Books.

Lear, J. (1998). 'The introduction of Eros'. In *Open-Minded* (pp. 123–147). Cambridge, MA/London: Harvard University Press.

Lear, J. (2003). 'The idea of a moral psychology: the impact of psychoanalysis on moral philosophy in Britain'. *The International Journal of Psychoanalysis 84*: 1351–1361. (p. 636)

Lear, J. (2005). 'Philosophy'. In E. Person et al. (eds), *Textbook of Psychoanalysis* (pp. 513–525). Washington, DC/London: American Psychiatric Publishing Inc.

Loewald, H. W. (1978). *Psychoanalysis and the History of the Individual*. New Haven, CT/London: Yale University Press.

Lomas, P. (1999). *Doing Good?* Oxford: Oxford University Press.

Mahler, M. (1986). 'On the first three subphases of the separation-individuation process'. In P. Buckley (ed.), *Essential Papers on Object Relations* (pp. 222–232). New York/London: New York University Press.

Mitchell, S. (1988). *Relational Concepts in Psychoanalysis*. Cambridge, MA/London: Harvard University Press.

Money-Kyrle, R. (1952). 'Psycho-analysis and ethics'. *The International Journal of Psychoanalysis 33*: 225–234.

Nietzsche, F. (1882/2001). *The Gay Science*, trans. J. Nauckhoff, ed. B. Williams. Cambridge: Cambridge University Press.

Plato (1961). *Republic*. In Plato, *The Collected Dialogues*, eds E. Hamilton and H. Cairns. Bollingen Foundation. Princeton, NJ: Princeton University Press.

Ridley, A. (1998). 'What is the meaning of aesthetic ideals?' In S. Kemal et al. (eds), *Nietzsche, Philosophy and the Arts* (pp. 128–147). Cambridge: Cambridge University Press.

Sagan, E. (1988). *Freud, Women and Morality*. New York: Basic Books.

Sherman, N. (1995). 'The moral perspective and the psychoanalytic quest'. *Journal of the American Academy of Psychoanalysis 23*(2): 223–241.

Swanton, C. (2003). *Virtue Ethics: A Pluralistic View*. Oxford: Oxford University Press.

Swanton, C. (2011). 'Nietzsche and the virtues of mature egoism'. In S. May (ed.), *Nietzsche's On the Genealogy of Morality: A Critical Guide* (pp. 285–308). Cambridge: Cambridge University Press.

Waddell, M. (1998). *Inside Lives*. London: Duckworth.

Wall, S. (2013). 'Perfectionism'. In G. Gaus and F. D'Agostino (eds), *The Routledge Companion to Social and Political Philosophy* (pp. 342–352). Abingdon/New York: Routledge.

Williams, B. (1985). *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press.

Williams, B. (2000). 'Naturalism and genealogy'. In E. Harcourt (ed.), *Morality, Reflection and Ideology* (pp. 148–161). Oxford: Oxford University Press.

Winnicott, D. W. (1967). 'The location of cultural experience'. *The International Journal of Psychoanalysis 48*: 368–372.

Winnicott, D. W. (1985). *Playing and Reality*. Harmondsworth: Penguin.

Winnicott, D. W. (1990). 'The theory of the parent–infant relationship'. In *The Maturational Processes and the Facilitating Environment*. London: Karnac.

Notes:

(^{1}) Largely for worse. Once upon a time, to work 'in virtue ethics' was simply to have an interest in the virtues—even though nobody who *was* interested in them called it that. Now the phrase can also mark adherence to an ethical theory in which the concept of virtue plays the same fundamental role as the concept of a good state of affairs in a consequentialist theory.

(^{2}) See also Waddell (1998: 3), for whom it's about 'the moral and emotional growth of the self, the character'; Lomas (1999), Harcourt (2015, 2018b), and Sherman (1995).

(^{3}) By 'psychoanalysis' I do not mean the practice of psychoanalysis but the theory or theories associated with it. Moreover I don't just mean the theory or theories associated with psychoanalysis as practised in the world's Institutes of Psychoanalysis, but also the cognate theories at large in the 'psychoanalytic diaspora' of broadly Freud-inspired psychodynamic psychotherapies.

(^{4}) Of course there has been another shift too, namely a tendency to not think of psychoanalysis as appropriate to the explanation or treatment of psychotic illness (Bateman and Holmes 1995: 213). But this point is neutral with respect to what I am arguing here: if you subtract psychosis throughout the history of psychoanalysis, what is left over now is seemingly very different from what was left over in Freud's day.

(^{6}) There is an odd gap, at this point, in Freud's theory of therapeutic efficacy: if what explains symptom-formation is the absence of a virtue, one would expect part of the theory of therapeutic efficacy to be a story about how the missing virtue is acquired, but Freud's therapeutic repertoire of free association, interpretation etc. does not seem to contain one.

(^{7}) [Eds: For further discussion of these ideas, see Gipps and Lacewing, 'Introduction: Know Thyself', this volume.]

(^{8}) A special case of work in Freud's thought, which it would take too long to go into fully here, is of course art. For Freud, the artist 'in a certain fashion actually becomes the hero, the king, the creator ... without following the long roundabout path of making real alterations in the world' (1958: 224). This characterization of the difference between the artist and the psychotic depends on Freud's theory of sublimation ('shifting the instinctual aims in such a way that they cannot come up against frustration from the external world', 1964b: 79). It is strange, though, that Freud didn't seem to see how much *real*—in addition to symbolic—change the artist needs to effect in the world, by way of transforming his or her various media, or therefore how many opportunities for frustration he or she encounters.

(^{9}) A striking example of this view in the contemporary 'virtue ethics' literature is Christine Swanton, for whom the difference between the virtues of 'mature egoism' and their defective cousins which superficially resemble them—'immature egoism' and 'self-sacrificing altruism'—consists in the presence or absence of certain 'depth-psychological traits' (2003: 10; 2011: 288).

(^{10}) I do not mean to imply that this is the *only* respect in which psychoanalysis might contribute to 'virtue ethics'.

(^{11}) 'With complete temporal balance and harmony between the various [developmental] lines, the result could not fail to be a completely harmonious, well-balanced personality' (A. Freud 1998: 139).

Edward Harcourt

Faculty of Philosophy, University of Oxford, UK

Oxford Handbooks Online

Abstract and Keywords

What has gone wrong with humanity that it is the only species that is destroying its habitat and justifying it as good or inevitable? The quality of a child's early nest, including experience with mother and others, co-construct the nature of the child, shaping a highly intertwined biology and sociality. The species-typical nest that humans inherited for their young facilitates the development of a self that is relationally attuned and communal and receptively intelligent. Neurobiological studies today support the general insight from psychoanalytic theory that early experiences with caregivers initialize the formation of the self. But civilization typically does not provide the evolved nest, promoting a species-atypical pathway for social and moral development, one of self-focused protectionism. Cultures that adults build and perpetuate are founded on the human dispositions brought about by early experience.

Keywords: evolution, evolved nest, evolved nest, small-band hunter-gatherer, moral sense, self, neurobiology, social schemas, toxic stress, chronic stress

Darcia Narvaez

(p. 637) Introduction: The 'Evolved Development Niche' as a Baseline for Human Well-being

WHAT has gone wrong with humanity?[1] How can it be that a particular human culture[2] has come to dominate most of the earth, acting in effect to destroy the planet (e.g. Millennium Ecosystem Assessment 2005; Ceballos, Erlich, and Dirzo 2017) and at the same time creating a cultural narrative that justifies such destruction (Ferguson 2012; Korten 2015)?

Freud (1929/2002), too, wondered about the state and behaviour of humanity and drew generalizations from his patients. His work led him to hypothesize an inborn id as a primary source of difficulty. Contrasting with Freud's nativism, Hillman and Ventura (1992) fingered a century of psychotherapy for inventing narratives for their patients on which to blame what they viewed as socially caused illnesses. In interviews by Ventura transcribed in the book, Hillman points out that other societies had less egocentric and isolating forms of dealing with life's troubles; they would never blame childhood experience for them, as psychotherapy can tend to do.

However, Hillman assumed that childhoods were similar across cultures (as, more recently, does anthropologist David Lancy 2015). This is untrue. This chapter will argue, against Freud, that his conception of the id represents a misshapen human nature that (p. 638) emerges from the denial of evolved needs through an abandonment of the 'evolved nest', common in 'civilized' societies. And it will argue, against Hillman, that the dominant culture (industrialized capitalism) has shifted away from providing children with what they evolved to need, perhaps especially in the USA whose overall health continues to worsen (e.g. National Research Council 2013) and which exports its child-raising ways to other nations.

One of the problems with many assessments of human well-being is a lack of a baseline against which accounts of normal and deviant development may be judged. Scholars often assume their own experience of contemporary life as a baseline (Pauly 1995), leading to mistaken assumptions and interpretations. For example, in a widely cited study used to support parental detachment, Robert LeVine and colleagues (LeVine et al. 2008) compared the infant caregiving of Boston mothers with that of the Gusii in East Africa. They used as a baseline for normal human behaviour the Boston mothers' high verbal interaction with their infants, assuming it was necessary for successful child outcomes. They were surprised that Gusii adolescents showed no signs of deprivation despite the fact their mothers were not verbally 'attentive' when they were young. Actually, the researchers were using the wrong baseline for normal human infant care. The Gusii infants and children were cared for in the manner that characterized 99 per cent of humanity's genus history (Konner 2005), with non-verbal responsiveness and extensive affectionate touch that kept the infants calm and growing well—as social mammalian young evolved to need, and which are known now to play a large role in the development of neurobiological structures and socioemotional intelligence (Narvaez 2014; Narvaez, Panksepp, Schore, and Gleason 2013; Narvaez, Wang et al. 2013). Unfortunately, it is common for research to be based on faulty assumptions, leading to faulty conclusions (Henrich, Heine, and Norenzayan 2010). Establishing baselines from the outset is a way to avoid the creeping biases that otherwise enter into conceptualizations of research and its interpretation.

A first baseline that should be mentioned is the highly cooperative nature of the natural world. The biosystems of the planet are all cooperative, operating in a gift economy of give and take and give (Bronstein 2015; Margulis 1998; Paracer and Ahmadjian 2000). Competition is a thin icing on the thick cake of cooperation. In fact, each human being is

a community of cooperation with trillions of microorganisms that keep alive the individual, who carries 90–99 per cent non-human genetic material (Dunn 2011). Only recently in human genus history has a dominator culture emphasized a competitor model that undermines the cycle of give and take in the natural world and represents a shifted baseline about how the world works (Worster 1994). Even within the family, cooperation is fundamental where costs and benefits balance among members over time.

Second, we need a baseline for human development against which comparisons are made. The baseline I use is humanity's species-typical 'nest', niche, or developmental system (Gottlieb 2002). What is species-typical for a human being? Time-tested over millions of years by our ancestors, the human 'evolved developmental niche' or EDN (Narvaez 2016) represents the inherited characteristics of early life experience provided for 99 per cent of human genus existence (until agriculture came into existence, after (p. 639) which many societies decreased the provision of the EDN). Most characteristics of the human EDN emerged with social mammals over thirty million years ago and intensified with human evolution—(a.k.a. 'hunter-gatherer childhood model'; Konner 2005). The EDN for young children includes: soothing birth experience (no separation of mother and baby, no inducement of pain, e.g. smacking, suction, or skin pricks, which is routine today); responsiveness to needs (preventing distress); infant-initiated breastfeeding for several years; extensive positive touch (nearly constant in first year) and no negative touch; multiple, responsive adult caregivers and maternal support; social climate that promotes positive emotions; free play in nature with multiple-aged playmates (throughout childhood and beyond) (Hewlett and Lamb 2005). All these components optimize normal development both neurobiologically and psychologically (for reviews, see Narvaez, Panksepp, Schore, and Gleason 2013; Narvaez, Valentino Fuentes, McKenna, Gray 2014). Why is the EDN important for optimizing normal development? Because the human neonate is born at least eighteen months early compared to other primates, with only 25 per cent of the adult-sized brain at full-term birth, so most brain and body systems develop and form postnatally (Trevathan 2011).

The quality of the early nest, including experience with mother and others, co-constructs the nature of the child, shaping a highly intertwined biology and sociality. In fact, humans are biosocial creatures through and through. That is, the functioning of a child's neurobiology is highly influenced by social experience and the biology that develops from that experience guides the nature of the child's sociality. Significant foundations are established during sensitive periods in the early years across brain and body systems, and other capacities are layered upon these. Life is an ongoing biosocial experience wherein gene expression shifts from experience to experience, continuing to build physiologically based habits of getting along in life.

Neurobiological studies today support the general insight from psychoanalytic theory that early experiences with caregivers initialize the formation of the self (Panksepp and Biven 2011).[3] The self's emotionally dynamic structure, with elements that importantly are non-conscious, derives from biosocial learning in childhood. In what follows, I examine how the self relies on the development of particular embodied knowledge structures or

schemas through interaction with caregivers, which psychoanalysts have theorized as object relations. But it is important to realize that early co-construction of the self can go in a species-*typical* or a species-*atypical* manner. Too many selves in existence today have been initialized in a species-*atypical* nest, leading to disordered emotional foundations that build ethical orientations of self-protectionism instead of communalism.[4]

(p. 640) In what follows, I explore and explain the importance of the EDN for our understanding of the social and moral self, providing a wider empirical context for psychoanalytic ideas concerning the centrality of object relations in child development. I begin with the idea of social schemas.

Social and Self Relational Schemas

The species-typical nest that humans inherited for their young facilitates the development of a self that is relationally attuned and communal. For a compassionate ethical self to be fostered and maintained throughout childhood, adolescence, and early adulthood, particular learning experiences are needed. Even in adulthood, attention to the nature of our environments and experiences is required. In our small-band hunter-gatherer past (our 99 per cent), this took place as a matter of course (Narvaez 2013, 2015).[5] The focus here will be on early childhood experience.

Starting perinatally, the sense of self is an emergent property, an outcome of biosocial experience, based in the development of implicit socioemotional intelligence that includes a deep attunement to the social world.[6] The moral self is formed through relational experience. In early life, reciprocal communication and companionship care convey to the infant that s/he matters, that the feelings the infant communicates matter, that relational communion is the normal way to be in the world. From beginnings in the womb and under good care after birth, for the child the world of experience is centred around Mother, gradually expanding to include a community of other 'mothers'. The gradual distinction of the internal and external environments occurs through experience with multiple caregivers who sensitively bring to life the child's self within a social world.[7]

Experiences in early life establish how and how well the brain develops, not only for health and intelligence but sociality and morality.[8] It is not surprising that early experience has long-term effects on all biological systems, as humans are complex dynamic systems. Interacting with the maturation and growing history of the child, early experience co-constructs the cognition–emotion–action patterns that form the implicit world view the child takes along throughout life. One might say that early life furnishes apprenticeship for social, cognitive, and moral development.

From a cognitive science perspective, general understandings of how the world works form from the accumulation and transformation of specific experiences. In cognitive (p. 641) information processing theory, these general understandings are often called schemas (Piaget and Inhelder 1969; Taylor and Crocker 1981). As generalized knowledge structures residing in long-term memory, schemas organize an individual's operational

activities (Piaget 1970; Rummelhart 1980; Taylor and Crocker 1981). Schemas develop from repeated experience, drawing generalizations from recurrent patterns. Their form is tightly organized but flexible in accessibility and adaptation—each instantiation of a schema alters the schema. More than 'in the mind', they connect emotion, cognition, and behaviour, subsuming procedural knowledge (knowing how), declarative knowledge (knowing that), as well as embodied cognition formed from situated (context-specific) experience, corresponding to ways of being. Schemas arise from repeated opportunities to solve a particular problem; since humans face similar problems, many schemas are similar across individuals (Marshall 1995). Some basic schemas developed in early life are related to the physical world (e.g. gravity's effects on objects we drop; Piaget 1954). Other early-forming schemas concern the social world, are more dynamic, and fluctuate with each situation. These dynamic forms are part of a set of schemas we develop for the self and the social life.

Social schemas reflect patterns of generalized experience applied within or across social situations or in ways of being for particular relationships. Some schemas are chronically employed by an individual (e.g. assessing others based on weight) and become 'chronically accessible'—easily and quickly deployed because of frequent use (Mischel and Shoda 1995). These can be anchored to particular situations (e.g. moments with mother) or generalized across them (relations with women). Chronically accessible schemas influence how an individual processes life events, directing attention, filtering and organizing stimuli; these schemas influence the selection of life tasks and goals and translate into behavioural routines that become highly practised and automatic (Cantor 1990). Thus, chronically accessible schemas contribute to the form of a personality, varying by particular situation. These activated schemas guide expectations, winnowing and chunking experience, and shape habitual reactions. They include as a subset the schemas that psychoanalysts have termed 'object relations' (rather unfortunately, given that such schemas concern relations between people/subjects).[9]

Schemas and Object Relations

On Westen's (1990: 686) definition, object relations are 'the cognitive, affective and motivational processes mediating interpersonal functioning, and … the enduring patterns of interpersonal behaviour that draw upon these intrapsychic structures and processes'. Object-relations theory argues that the need for relationships with others is primary, rather than dependent—as Freud thought—on the satisfaction of other desires (p. 642) (for nourishment or bodily pleasure). Interpersonal relations are the framework within which child development, especially emotional and social development, takes place.

Central to this is the thought that the self is built up from representations of significant others and of relationships to them. These include not only identifications the child makes with others, but also experiences of the self in relation with others. Thus, in her self-formation, the child constructs psychological ('internal') models of herself, of others, of particular relationships with others, and of relationship structures more generally. Not

only are relations to *actual* others of central importance to the child's developing social and moral self, so are the child's *experiences* of such relations and *interpretations* of those experiences. Thus, while actual events in childhood impact on the formation of a child's object relations, this impact is mediated by the child's developmental level, prior experience, and an already-emerging personality structure.

The similarity between object relations and the concept of social and self schemas is clear from the descriptions provided, and Westen makes the link explicit (1990: 687). He argues that object relations are best understood as organized networks of associations and encoded in many different modes of representation (including linguistic, imagistic, bodily, and procedural knowledge; here I add neurobiological). One finds similar connections drawn in the conceptual frameworks of other theories that seek to straddle the empirical and psychoanalytic models, including the 'internal working model' of attachment theory (Cassidy and Shaver 1999), Stern's (1985) concept of 'RIGs' (Representations of Interactions that have been Generalized), and Beebe and Lachmann's (2002) 'interactional structures'. As Eagle (2013: 887) comments, although there are differences between these theories of unconscious self–other representations, they have much in common, in particular that such representations are acquired early in childhood as a result of repeated patterns of interaction with others, that they involve implicit and procedural knowledge, and that they strongly influence one's relations with significant others later in life. Understanding object relations as schemas does not detract from nor overlook the significance of the specifically psychodynamic elements of object relations, such as their role in psychic conflict and transference (see e.g. Andersen and Thorpe 2009; Westen and Gabbard 2002a, 2002b). Rather, the theory of self and social schemas can be enriched by attention to their psychodynamic operations.

Early Social Schemas and 'Vitality Dynamics'

Thus, according to object relations and social schema theories, understandings of the social world are shaped by early experience, specifically by the mothering (nurturing responsive care) received from mothers, fathers, and others. A vibrant, true self develops within relationships of mutuality that 'affirm, validate, acknowledge, know, accept, understand, empathize, take in, tolerate, appreciate, see, identify with, find (p. 643) familiar … love … '—what Jessica Benjamin calls practices of *mutual recognition*, experiences commonly noted in research on mother–infant interaction: 'emotional attunement, mutual influence, affective mutuality, sharing state of mind' (Benjamin 1988: 15–16). Simultaneously moulding neurobiological and psychological structures, these experiences ground expectations for the social relationships to come.

Humans naturally experience others in terms of vitality or energy dynamics (Stern 2010). They evaluate others in terms of their emotions and mental states, what they might be thinking and what they really mean, their authenticity in self-expression, anticipating what they are likely to do next, as well as the status of their health, on the basis of the fluctuating vitality expressed in their nearly constant movement. At least five dynamic

properties are linked together as a gestalt in our (and babies') experience of vitality: time, movement, intention/directionality, force, and space. These are apprehended together without conscious awareness (though may be studied in isolation).

Vitality dynamics become chronically accessible schemas for ongoing social behaviour. The early social life of the baby shapes capacities for recognizing and performing holistic dynamic movement in the social world through experiences in everyday life: the practice of micro-communicative movements with caregivers. From two and a half months to about six months of age, the baby's behaviour is highly responsive to parental behaviour through touch, facial expressions, and vocalizations. In the Western world, 'face-to-face play becomes the main "game" ' (Stern 2010: 106). Generally, some culturally appreciated combination of communications take centre stage: eye gaze, facial expression, vocalizations, gestures, body tonus. Subsequently, as babies become better able to control movements, they easily mirror social partners and initiate social events. Interactions become rich and mutual.

Daniel Stern asks why nature 'planned' for babies not to speak or understand words for the first year or so of life. His answer is that they 'have too much to learn about the basic processes and structures of interpersonal exchange', such as 'the forms of dynamic flow that carry social behaviour' which they must learn 'before language arrives to mess it all up' (Stern 2010: 110). The basic structures of interpersonal exchange 'are all non-verbal, analogic, dynamic Gestalts that are not compatible with the discontinuous, digital, categorical nature of words' (Stern 2010: 110). What kinds of things is a baby prepared to learn socially in the first eighteen months?

- Mutual eye gaze practices (how, how long, with whom)
- How to read postures
- How to solicit others for needs such as food or play
- 'Rules' of games
- Turn taking
- Greetings
- Joking
- Expressing affection
- Making friends
- And much more

(p. 644) This implicit, relational, and non-verbal knowing that is predominately processed in the right hemisphere of the brain rapidly develops, under conditions of good care, in the first years of life. Parent 'matching/mismatch of vitality forms can shape what the infant does and how he feels about doing it. It is like sculpting his mind from the inside out' (Stern 2010: 115). The interpersonal dynamics of synchronizing motives, intentional states, and behaviours with another—the forming of a duet of 'being with' the other person by participating in the dynamic flow—can be described as a *communicative*

musicality (Trevarthen 1999). Each relationship is a way to implicitly learn an alternative pattern of social being: 'Dad and I relate this way; Mum and I relate that way'. The child learns to be alive together with each caregiver, an aliveness missing in many who suffer early trauma (Clancier and Kalmanovitch 1984). Generalizations about relationships are made from these early experiences.

Human infants, like other apes, develop strong attachments to caregivers in the first year of life. In fact, the neurobiology of attachment appears to ground lifetime brain function as well as social and moral behaviour (Gross 2007). Emotions that guide 'attachment for companionship' are innate and test possibilities and values of shared experience (Trevarthen 2005). Loving care is conveyed by the caregiver's enjoyment of bathing, clothing, and feeding the child—'it's like the sun coming out, for the baby' (Winnicott 1957: 14–15).

Psychologically, the ongoing (emotional and physical) support that caregivers provide communicates to the young child the trustworthiness of his body signals and the safety and supportiveness of the world in getting needs met. Consistent responsiveness leads to a self highly secure and deeply rooted in the social landscape, a self who skilfully derives pleasure from and prosocially contributes to the community. When caregivers are not ongoingly supportive (e.g. isolating the baby from touch and calming comfort), the child's foundational neurobiology and sense of living can form instead around a sense of danger (Sandler 1960), along with a sense of rejection or negation (Litowitz 1998).

Sandler (1960) suggested that the early sense of danger grows into cynicism or anxiety, minimally into an adult with little trust or confidence in the self and the world, and maximally into an adult with personality disorders. The insecure self harbours a sense of abandonment and badness, apparent in insecure attachment, which unconsciously flavours interpretation of life experience and propels behaviours to avoid those feelings, demonstrated in neurobiological inflexibility ('stiffness' of the mind or 'heart'). Using the terms of the early object-relations theorists, such as Klein and Fairbairn, we may seek to understand this as involving the internalization of a 'bad object', although such phrasing is heavy with various controversial theoretical commitments, especially if internalization is understood as a defensive process motivated by anxiety, rather than the implicit development of a schema through habituation.

But however one further theorizes the process, we may agree that early dynamic experiences undergird our expectations and sensibilities for social life, rooted in neurobiologically grounded 'narratives' or schemas for the self: 'I am good and competent and the world is to be trusted' versus 'My urges are bad and the world is to be (p. 645) distrusted' (Narvaez 2011). 'Dynamic forms of vitality are part of episodic memories and give life to the narratives we create about our lives' (Stern 2010: 11).

When a young child's evolved needs are met through companionship care, including experiences of ongoing intersubjectivity with familiar, loving others, cooperation with self and others becomes an intuitive baseline for life (Narvaez 2014; Trevarthen 2005). Are

there other early caregiving practices that are critical to the formation of a social and thence a moral self?

Species-Typical Child-Raising and the Evolved Developmental Niche

A key oversight of Freud that still continues within most branches of the social sciences is the failure to understand the importance of attending to the particular types of early care a child experiences for shaping specific individual capacities, which Stern (2010), for example, tried to explicate. (Researchers who are starting to focus on species-expectant care include Greenough, Black, and Wallace (1987) and McLaughlin, Sheridan, and Nelson (2017).) Winnicott (1957) emphasized the 'good enough' holding and handling environment that allows a baby to feel omnipotent until ready to face the reality principle. But we can go further and be more specific if we consider *typical* and *atypical* child-raising from an evolutionary perspective. Species-typical child-raising involves the evolved developmental niche (EDN) described in the Introduction (soothing perinatal experience, extensive breastfeeding and affectionate touch, quick and warm responsiveness to needs, multiple responsive adult allomothers, positive social support and climate, self-directed free play). These components provide the environment for optimal development, matching up with and constantly interacting with the maturational development of the child with epigenetic and plasticity effects. *Atypical* child-raising occurs outside the nest, much like Harry Harlow's (1958) monkeys, raised without maternal love, touch, responsiveness, and breastfeeding. The results are disturbed development.

Most people, including scholars and practitioners, still lack awareness of the type of child-raising that is typical for our species and how strange and atypical civilized child-raising is for the human species (although I think that perhaps many psychotherapists suspect that something went wrong with the childhoods of their patients, that mistreatment or trauma occurred at some point from babyhood to early adulthood). Early life experience may be the most critical factor in shaping the type of human being that is raised. Schemas for self in relation to other, representative of the quality of one's neurobiological self-regulation and social growth, can go awry in a variety of ways depending on the nature of undercare.

As noted by Winnicott, the ecology of the child is a system: 'The unit is not the individual, the unit is an environment-individual set-up. The centre of gravity of the being (p. 646) does not start off in the individual. It is in the total set-up' (Winnicott 1975: 99). Along similar lines, Urie Bronfenbrenner's (1979) ecological systems theory identifies several levels of interactive influence in the child's life (e.g. family and peer relations, relations of the family with the community, job pressures on parents, cultural and historic events). Bronfenbrenner's theory provides a *vertical* view of a child's life which is complemented by the *horizontal* view of development across generations provided by evolutionary systems theory (Oyama 1985). This focuses on a host of evolutionary inheritances (genes

Subscriber: UC - Los Angeles (UCLA); date: 17 March 2019

are only one of many inheritances) such as the developmental system or nest, maternal ecology (health, nutrition), and the cultural and environmental ecologies left by prior generations, among other extra-genetic features. Relational developmental systems meta-theory (Overton 2013, 2015) focuses on the *ongoing interactions* of individual and context as the individual self-organizes through developmental processes. Each theory points to the *constant dynamic interplay* of many elements in the developmental life course. Ethogenetic theory (Narvaez 2018) integrates these theories to describe the nature and pathways of humanity's moral development, focusing on inheritances and neurobiological foundations that are epigenetically and plastically integrated in early childhood as the child functionally adapts to the social and physical environment.[10]

The child has a built-in maturational schedule, in which layers of foundation are laid for more complex, later-developing systems, all self-organizing during sensitive periods (Knudsen 2004). Caregiver interaction with a child is not a moulding of clay but the layering of a rapidly growing agent. Physiologically, multiple systems are setting parameters and thresholds based on maturational schedule and the nature of early care. Of course, proper environmental 'holding' ideally is provided by adults who themselves were well cared for, who are willing to adapt to the needs of the baby in the critical first months and years.

One complex and fundamental set of capacities highly influenced by early experience is basic self-regulation, involving regulatory systems that are all too often considered innate and involuntary (e.g. the stress response).[11] We know now that most self-regulatory systems (e.g. the stress response) are shaped by experience during sensitive periods inside and outside the womb (Lupien, McEwen, Gunner, and Heim 2009). Multiple neurobiological systems and layers of systems (e.g. those underpinning other systems such as vagal tone, Porges 2011) which become increasingly self-organized are nurtured into being by postnatal care. Neocortical self-control of arousal (through the later-developing cortex) is not in place till the end of the first year, so maintaining proper arousal is largely up to the caregiver, who keeps baby from over- or under-arousal. (p. 647) Of course, the baby has limited capacities for self-calming (although can do simple things like turn away or close eyes), but otherwise relies on caregivers to guide the development of self-regulatory capacities.

Babies are optimally aroused to play and be happy in a particular zone of arousal. Too little or too much parental stimulation shifts the baby out of the optimal zone. Caregivers help baby stay calm with skin-to-skin carrying (a remarkably calming tool), rocking and patting, and other touch-plus-movement behaviours. Representatively, the orbitofrontal system connects directly to the autonomic system and, when properly functioning, regulates its two subsystems (sympathetic and parasympathetic). The sympathetic system mobilizes the body for action (flight–fight) whereas the parasympathetic system conserves energy (freeze–faint) to preserve life. When the orbitofrontal system is immature or underdeveloped from lack of comforting touch and responsive care, both the sympathetic and parasympathetic systems can mis-perform. The child will have difficulty adapting appropriately to situations and will show a deficit of empathy (Schore 2003a).

Instead of automatic self-soothing and self-regulation after a stressor, the individual will respond cacostatically—i.e. too strongly or too weakly, taking extra time to return to homeostasis.

Much of the social self is grounded in the early months and years when foundations for multiple systems are laid down, interactively shaped among themselves, and tuned up by experience. The same, I argue later, can be said for the moral self. Morality too emerges from neurobiologically wired capacities significantly shaped during this period. Before turning to that argument, it is worth briefly noting the species-typical context for human development provided by small-band hunter-gatherer communities.

Small-Band Hunter-Gatherers: Humanity's 99 Per Cent

Archaeological studies have indicated that the human genus has been around for over two million years.[12] And prior to the emergence of agricultural settled societies in the last 12,000 years or so (humanity's 1 per cent), all human beings spent their lives in small-band hunter-gatherer communities, hundreds of which still exist today. Anthropologists note the similarities of these communities which developed independently all over the world.

Small-band hunter-gatherers (SBHG, aka nomadic foragers) are groups of about five to twenty-five individuals whose membership shifts from day to day. They have no possessions, forming 'immediate-return' societies—consuming food resources as they are encountered and not hoarding resources, including cultivation of plants or domestication (p. 648) of animals. Despite living in vastly different landscapes around the world, they share similar cultural and personality characteristics. They live within a set of landscapes over which they migrate regularly. They live in highly cooperative but fluid groupings with deep values of egalitarianism and sharing. There is no coercion, even between adults and children, but high autonomy (Boehm 1999). All over the world SBHG provide the evolved developmental niche (EDN), which converging evidence suggests leads to the adult personalities that are also similar around the world: generous, calm, content, humble, kind, highly autonomous, and highly communal. Unlike the characters in the mythologies about them promoted by some scholars, SBHG are not warlike (they have no possessions to fight about), though jealous male rages are recorded.[13]

The 'Moral Sense'

For the remainder of this chapter, we turn to the question of the implications of this broader view of the human situation in terms of the EDN for our understanding of morality and the place of social and self-schemas/object relations in moral development.

First, what do species-typical experiences bring about in terms of morality? Species-typical upbringing leads to the moral capacities noted among the adults of indigenous peoples around the world by first-contact explorers and later by anthropologists, but perhaps most clearly by Charles Darwin.

In response to Herbert Spencer's early arguments that selfishness and aggression were part of human nature, Darwin (1871) brought forth the idea of an evolved 'moral sense'. He identified its characteristics—social pleasure, empathy, concern for the opinion of others, and habit development for the sake of the community—and how they were apparent in other animals that had evolved prior to human beings, making the moral sense part of human nature, not the other way around. He assumed these were inherited characteristics. But recent evidence suggests otherwise—that they develop after birth and may require the early nest (Kochanska 2002).

Ethnogenetic theory describes the early-life neurobiological foundations of an individual's preferred moral orientations (Narvaez 2016). For example, normal, species-typical, human development, with EDN provided and secure attachment established, fosters a relationally attuned, emotionally present, heart-centred orientation. How do we know this? We can see what is species-typical among small-band hunter-gatherer communities who provide the evolved developmental niche for children throughout their lives and end up with the type of personalities described above (for reviews, see (p. 649) Ingold 2005; Narvaez 2013). (The alternative, self-centred orientations that develop with species-atypical care are described later in the section on self-protectionism.)

The key to human development, and especially social and moral development, is relationship, beginning with conception, gestational experience, and early life. Many babies died when it was assumed that physical care (food, warmth, nourishment) was enough (Spitz 1945). As emphasized previously, early relationships provide the soil for the growth of neurobiological systems (e.g. neurotransmitter function and proliferation, stress response system function, vagus nerve myelination). During this time much more than internal working models or object relations are set up: the multiple neurobiological systems set thresholds and parameters for life. Our psyches rely on these psychobiosocial foundations.

Firm foundations for a good life include the proper shaping of the visceral-emotional nervous system on the hypothalamic-limbic axis (Panksepp 1998), which includes multiple limbic and subcortical structures and multiple types of neurotransmitters. Evolved emotion systems of *care* and *play* (the italics denote physiologically mapped emotion systems), which guide sociality and learning (Konner 2002), are the primary systems for the social self, self-identity, and sense of what is true (MacLean 1990). 'Knowledge of the world is gained by moving about in it, exploring it, attending to it, ever alert to the signs by which it is revealed' (Ingold 2011: 55). So too, the social world. From our embodied experience, we learn to see and to be. Our knowledge is situated in the direct engagement with the world. Existing in a dynamic flow of being is enlivening. ' "Forgetting" ourselves is how we lose our sense of separation and realize that we are

not other than the world' (Loy 2002: 7). This communal flow is learned in the early nest which engraves expectations for relations with others, i.e. sets up the child's social schemas. (When this goes well then, as object-relations theory has it, the child has 'internalized a good object'.)

Human brains are designed to be reward-seeking, primarily from social relationships (Panksepp and Biven 2011). However, these inclinations are shaped postnatally. With immersion in supportive and socially stimulating family and community, the individual will grow the capacities for sociality—micro-skills for getting along well, along with pleasure in being with others because it is so much fun. Empathic and sympathic capabilities will grow as a matter of course. If the individual is minimally supported socially and spends most time alone or away from intimate relationships (as most USA children do today), these capacities will not grow properly. The child will have fewer social and emotional capacities and find social relations less pleasurable as a result. Deficiencies will be mappable in endocrine and neuronal structures (e.g. Pollak and Perry 2005; Weaver, Szyf, and Meaney 2002). Social engagement and compassionate functioning will be impaired. The individual likely will have greater hostility and distrust towards others (Kruesi, Hibbs, Zahn, Keysor, Hamburger, Bartko, and Rapoport 1992).

With companionship care, capabilities for emotional presence and secure attachment develop that allow for an *engagement* ethic, an orientation of egalitarian relational (p. 650) attunement.[14] Shaped initially by early learning of an intuitive dance in relationships with caregivers (as described earlier), and then from envelopment in a broad 'circle of attachments', the child simultaneously practises and grows capacities for reverence, intersubjectivity, play, and egalitarian respect (Narvaez 2014). The formation of a positive sense of self and relationships was well described by R. D. Laing who observed that ontologically secure persons:

> have a sense of … presence in the world as a real, alive, whole, and in a temporal sense, a continuous person. As such [they] can live out in the world and meet others: a world and others experienced as equally real, alive, whole, and continuous.(1959/1990: 39)

Regular experiences of mutual recognition lead to a sense of deep relatedness and being known.

A species-typical nest promotes the development of *losing the self* initially in caregivers' arms and then in social pleasure, in communal trance or dance. Because of the early formation of secure attachment through companionship care, deep trust of the social world is foundational; socially attuned relating to others forms the framework for a good life. The deep, routine experiences of empathy, resonance, intersubjectivity with caregivers 'tunes up' the child's neurobiological self (e.g. vagus nerve, internal working models of how the world works) to at first dynamically mimic and then become what is experienced. What the young child practices is what she becomes. Routine experiences of empathy, social attunement, and social pleasure lead to values of compassion, social harmony, and togetherness as a matter of course. I–Thou, rather than I–It, relations

become the baseline for social functioning. The life course and personality are set to be other-regarding rather than self-centred (Narvaez 2014).

Although an engagement ethic orientation is sufficient for many situations, humanity's fullest moral flowering occurs with the full development of the neocortex and related thalamic structures, a somatic-cognitive nervous system on the thalamic-neocortical axis (MacLean 1990; Panksepp 1998). These systems take several decades to complete themselves with ongoing supportive experience (three decades for executive functions; four or five decades for synthesizing broadly). Higher capacities for executive functions (e.g. foresight, goal-coordination) and reasoning continue to develop into adulthood; moral reasoning too develops more sophistication with maturation and experience (Rest, Narvaez, Bebeau, and Thoma 1999). These capacities can be rooted in the sensorimotor social experiences of early life and develop gradually. Within a species-typical developmental system, executive functions are grounded in communality, curtailing their use in a manner contrary to well-being of the community. (In a species-*a*typical developmental system, executive functions may develop in a more self-centred direction from missing (p. 651) the companionship care that is provided in a species-typical system. Moral reasoning can then involve only intellect and operate in an emotionally detached manner.)

Communal imagination is rooted in the engagement ethic, but adds abstraction and metacognitive capabilities (Narvaez 2014). Communal imagination coordinates instincts, intuitions, and deliberation, as well as internal and external information, for intentional compassionate action that attends to steps, obstacles, and consequences. Imagination is integrated with empathy and attachment. That is, it is far-seeing but not far-hearted. It is grounded in concern for one's particular community but also for communal welfare in distant times and places, for those not present in time or place. Societies where children are raised within a species-typical development system also support the development of wise elders. Wise elders display communal imagination capacities to the greatest extent, and some societies also build it into their planning across generations (e.g. Japan; Native American peoples with concern for the seventh generation).

Self-Protectionism

Civilization presents an alternative pathway for social and moral development, one of self-protectionism. What does the self-protective pathway look like? Let's suppose a child who experienced some degree of a species-atypical early life: traumatized by painful procedures at birth, left largely alone in carriers, playpens through the day, and cots at night, who is left to cry alone because parents are told this is best or even want their own independence. These are examples of a violated evolved nest, what I call 'undercare'. Undercare shifts energies away from growing the emotion systems of sociality (care, play), scheduled to develop after birth, and instead enhances the survival systems with which we are born—emotion systems of *fear, anger, panic/grief, seeking* (anticipatory

euphoria that guides exploratory and acquisitive behaviour) and primitive lust (Panksepp 1998).

An early life of inadequate or under-care results in layers of deficiencies, from brain structures to hormonal regulation to structures underlying the psyche and integral to overall system integration. Such an individual has not been able to develop expectations and internalization of good-enough others who relate to the individual in supportive, positive ways. Of course, the problems go 'all the way down' to endocrine and neurotransmitter function, neural synapses and networks, paralleling impaired emotional self-regulation from the lack of early physiological shaping by responsive care during sensitive periods. These multilayered, somatosensory 'knowings' lead to the unique combination of experiential traces and fantasy, distorting the individual's view of the social world.

Set on the wrong trajectory, the child will not properly develop the neurobiological structures that are scheduled to develop during early life and that undergird species-typical (p. 652) sociality. When adults leave a child to cry in rage or panic, they are lubricating a dispositional orientation towards anxiety. In a dissociated state (detachment from the immediate situation), the individual is cut off from external and internal stimuli and fantasy takes over (Schore 2003a). A false, grasping self takes over, signalling a shift to the primitive survival systems of rage or ritual, false selves or false narratives that keep the self calm enough.[15] Instead of agile relational capacities, the individual stands apart. What becomes enlivening in species-atypical nests is often destructive of self or other.

When the stress response kicks in, self-protectionism predominates. As studies following terror-management theory demonstrate, activating a fear response lessens empathy and concern for others and increases distrust of outgroup members (Nisbett and Cohen 1996; Rosenblatt, Greenberg, Solomon, Pyszczynski, and Lyon 1989). Self-protectionism is characterized by an emphasis on survival mechanisms such as following precedent and maintaining territoriality (MacLean 1990). In this mindset, ruthless use of shaming, threat, and deception to control others seems like the right thing to do (Shaver and Mikulincer 2007; Staub 1992). When threat seems pervasive, strong-armed and strongman tactics seem reasonable.

Toxic stress in early life promotes a generalized threat reactivity, shifting one to conditioned responses (what worked for self-calming in the past), thwarting free will in the moment. Capacities for other-regarding ethics evaporate. The social world is persecutory. The individual sets up a sense of security around a harsh social world, of one kind or another, so as not to expect too much and thereby not be devastated again. Devastation can occur through such things as engulfment or petrification by the other or by personal implosion (Laing 1959/1990). Without a witness or advocate to process the lack of a proper early nest and childhood the individual becomes stuck in denial. A baby abandoned and unloved develops bad object relationships and, in a consequent negative transference towards people generally, projects her bad objects into others, thus isolating

herself from creative social attunement. Brain-behavioural disorders (e.g. personality disorders) are a natural outcome. Deep ontological insecurity (Laing 1959/1990) leads to social self-protectionism. The individual may favour one or another or bounce among social self-protectionisms: externalizing—hostility and aggression towards others, seeking control and dominance; or, internalizing—social withdrawal, depression, and anxiety (Perry 2000; Schore 2003a, 2003b). Each dissolves prosociality. Avoidance of a sense of non-being, of shame, and of annihilation shapes actions and reactions, with negative transferences to others predominating. Thus, self-protectionist ethics reflect an enhancement of the survival systems through early conditioning while right-hemisphere lateralized self-regulatory and relational capacities are underdeveloped or shut down. Unable to stand negative feelings towards the self (e.g. guilt), the individual slides into bullying or being victimized as comfortable psychic locations. What becomes normalized is a role in a dominance hierarchy, either through aggressive action or through appeasement or withdrawal. The self-protective individual is not relaxed and open but braced against others.

(p. 653) The Impact on Moral Imagination

In species-atypical developmental systems, the moral imagination develops awry. When the child's needs are not met:

> there will be a prematuration of the system; his sense will awaken precociously and the child will use understanding as a defence system against traumatism … intellectual activity, may be a bad thing that leads at first to a false system, in the wrong direction, without belonging to the body. A certain type of mental functioning—such as memorizing or cataloguing impingements with a view to assimilation at later stages of development—may be a burden for the psyche-soma, or for the continuity of existence that constitutes the self. This type of mental functioning acts as a foreign body if it is associated with inadequate adaptation to the environment.
>
> (Clancier and Kalmanovitch 1987: 60)

In the species-atypical nest, experiences of physical and emotional isolation build in a lasting detachment and distrust rather than a deep sense of empathy and relational responsibility (*detached imagination*, Narvaez 2014, 2016). Morality becomes only an intellectual exercise about hypotheticals or logic problems where intellect can be used to rationalize anything (see also Backström, this volume). Relational embeddedness and responsibility are alien. When one has learned to be detached from emotions (without awareness or intention), imagination too becomes detached from a relational grounding, making it possible to dream up and justify activities, actions, and products that harm others.

With significant undercare, trauma, or abuse, executive functions and reasoning may root themselves in *vicious imagination*, based on a splitting of the self that leads to rejection and the scapegoating or ruthless control of others. Missing are the early foundations of sociality that otherwise guide imagination and abstraction. In a threat-filled world, abstracting capabilities also become misguided, building on self-protective mechanisms. When one has learned to be socially oppositional, violence and aggression seem morally logical.

Conclusion

What has gone wrong with humanity? My contention is that species-atypical child-raising seeds underdeveloped and even skewed human nature and morality. The resulting misdeveloped adults create societies that perpetuate the undercare of children, fostering communities of people with neurobiologically framed paranoiac perceptions of the relational world.

Prior scholars have pointed out the inadequacy of modern child-raising arrangements. Many commentators on USA culture have pointed to the adolescent nature of adult (p. 654) culture today. For example, Paul Shepard (1998) pointed to adolescence as a key time when traditional societies provided initiation ceremonies into adulthood. As noted earlier, I think the misdevelopment begins much earlier.

Although psychoanalysts from Freud onwards have long intuited the importance of early experience and developed theories of object relations, they have had neither the data nor the tools to provide empirical support and context for these insightful intuitions. This has changed. We have extensive animal studies showing causal neurobiological effects of experience. We can see similar findings in human brains through retrospective studies of maltreated children. Clinical studies show the damaging effects of trauma and maltreatment in early life, while empirical approaches in child development have produced extensive evidence favouring schema theory.

However, both psychoanalytic and empirical descriptions of normal development have lacked an appropriate baseline against which accounts of normal and deviant development may be judged. The baseline of the early nest suggests that undercaring for children impairs human cognitive, social, and moral capacities. In fact, perhaps psychiatric (including psychoanalytic) descriptions of pathologies might be considered an atlas of deficits resulting from a degraded evolved nest. I contend then that psychic and moral health go hand in hand (see Harcourt, this volume). Thus, we might also compile an atlas of *moral* deficits from a degraded evolved nest

Moreover, the deformities that result appear to be linked to the destructiveness not only towards self and others but toward the other-than-human—the rest of the planet's living entities. In fact, Western scholars for centuries and the field of psychology from the beginning turned away from partnering with the natural world (Kidner 2001; Merchant 2003), with Freud (1929/2002) even suggesting that for civilization to move forward,

humans needed to take up the attack on nature to force it under human will, guided by science. Clearly this combination of child undercare, disregard for the natural world, and narratives to justify those behaviours have contributed to the ecological catastrophe we face today.

Human beings are complex creatures whose capacities are developed mostly after birth, unlike virtually every other animal. To grow well, each baby needs a set of responsive caregivers to provide the nurturing care, the mothering, that babies need to grow well (Hrdy 2009). Early experiences powerfully influence neurobiology and personality and ought to match up with what children evolved to need. The biosocially constructed self relies on the evolved nest. And so does our morality. We can build receptive intelligence and communal imagination or shrink the circles of moral awareness and concern to the self by not providing what babies and young children evolved to need. Which shall it be?

References

Andersen, S. M. and Thorpe, J. S. (2009). 'An IF-THEN theory of personality: Significant others and the relational self'. *Journal of Research in Personality 43*: 163–170.

Beebe, B. and Lachmann, F. (2002). *Infant Research and Adult Treatment: Co-constructing Interactions*. Hillsdale, NJ: Analytic Press. (p. 655)

Benjamin, J. (1988). *Bonds of Love*. New York: Pantheon.

Boehm, C. (1999). *Hierarchy in the Forest: The Evolution of Egalitarian Behaviour*. Cambridge, MA: Harvard University Press.

Bowers, M. E. and Yehuda, R. (2016). 'Intergenerational transmission of stress in humans'. *Neuropsychopharmacology 41*: 232–244.

Bronfenbrenner, U. (1979). *The Ecology of Human Development*. Cambridge, MA: Harvard University Press.

Bronstein, J. L. (ed.) (2015). *Mutualism*. New York: Oxford University Press.

Cantor, N. (1990). 'From thought to behaviour: "Having" and "doing" in the study of personality and cognition'. *American Psychologist 45*: 735–750.

Cassidy, J. and Shaver, P. R. (1999). *Handbook of Attachment: Theory, Research and Clinical Applications*. New York: Guilford Press.

Ceballos, G., Ehrlich, P. R., and Dirzo, R. (2017). 'Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines'. *Proceedings of the National Academy of Sciences (PNAS)*, published ahead of print 10 July 2017. doi: 10.1073/pnas.1704949114

Clancier, A. and Kalmanovitch, J. (1984/1987). *Winnicott and Paradox: From Birth to Creation*, trans. Alan Sheridan. London/New York: Tavistock Publications.

Cushman, P. (1995). *Constructing the Self, Constructing America*. Reading, MA: Addison-Wesley.

Darwin, C. (1871/1981). *The Descent of Man*. Princeton, NJ: Princeton University Press.

Davis, E. P., Glynn, L. M., Schetter, C. D., Hobel, C., Chicz-Demet, A., and Sandman, C. A. (2007). 'Prenatal exposure to maternal depression and cortisol influences infant temperament'. *Journal of the American Academy of Child and Adolescent Psychiatry 46*: 737–746.

Dunn, R. (2011). *The Wild Life of our Bodies: Predators, Parasites, and Partners that Shape who We are Today*. New York: Harper.

Eagle, M. (2013). 'The implications of conceptual critiques and empirical research on unconscious processes for psychoanalytic theory'. *Psychoanalytic Review 100*(6): 881–917.

Ferguson, N. (2012). *Civilization: The West and the Rest*. New York: Penguin.

Freud, S. (1929/2002). *Civilization and its Discontents*. London: Penguin.

Fry, D. P. (2006). *The Human Potential for Peace: An Anthropological Challenge to assumptions about War and Violence*. New York: Oxford University Press.

Fry, D. P. (ed.) (2013). *War, Peace and Human Nature*. New York: Oxford University Press.

Gottlieb, G. (2002). 'On the epigenetic evolution of species-specific perception: The developmental manifold concept'. *Cognitive Development 17*: 1287–1300.

Greenberg, J. and Mitchell, S. (1983). *Object Relations in Psychoanalytic Theory*. Harvard, MA: Harvard University Press.

Greenough, W. T., Black, J. E., Wallace, C. S. (1987). 'Experience and brain development'. *Child Development 58*: 539–559.

Gross, J. J. (ed.) (2007). *Handbook of Emotion Regulation*. New York: Guilford.

Harlow, H. (1958). 'The nature of love'. *American Psychologist 13*: 673–685.

Henrich, J., Heine, S. J., and Norenzayan, A. (2010). 'The weirdest people in the world?' *Behavioral and Brain Sciences 33*(2–3): 61–83.

Hewlett, B. S. and Lamb, M. E. (2005). *Hunter-Gatherer Childhoods: Evolutionary, Developmental and Cultural Perspectives*. New Brunswick, NJ: Aldine.

Hillman, J. and Ventura, M. (1992). *We've had a Hundred Years of Psychotherapy—and the World's Getting Worse*. New York: HarperCollins.

Hobson, R. P. and Lee, A. (1999). 'Imitation and identification in autism'. *Journal of Child Psychology and Psychiatry 40*(4): 649–659. (p. 656)

Hrdy, S. (2009). *Mothers and Others: The Evolutionary Origins of Mutual Understanding*. Cambridge, MA: Belknap Press.

Ingold, T. (2005). 'On the social relations of the hunter-gatherer band'. In R. B. Lee, R. B. Daly, and R. Daly (eds), *The Cambridge Encyclopedia of Hunters and Gatherers* (pp. 399–410). New York: Cambridge University Press.

Ingold, T. (2011). *The Perception of the Environment: Essays on Livelihood, Dwelling and Skill*. London: Routledge.

Kidner, D. W. (2001). *Nature and Psyche: Radical Environmentalism and the Politics of Subjectivity*. Albany, NY: State University of New York Press.

Knudsen, E. I. (2004). 'Sensitive periods in the development of the brain and behaviour'. *Journal of Cognitive Neuroscience 16*(8): 1412–1425.

Kochanska, G. (2002). 'Committed compliance, moral self, and internalization: A mediational model'. *Developmental Psychology 38*: 339–351.

Kolbert, E. (2014). *The Sixth Extinction: An Unnatural History*. New York: Henry Holt.

Konner, M. (2002). *The Tangled Wing*. New York: Owl Books.

Konner, M. (2005). 'Hunter-gatherer infancy and childhood: The !Kung and others'. In B. Hewlett and M. Lamb (eds), *Hunter-Gatherer Childhoods: Evolutionary, Developmental and Cultural Perspectives* (pp. 19–64). New Brunswick, NJ: Transaction.

Konner, M. (2010). *The Evolution of Childhood*. Cambridge, MA: Belknap Press.

Korten, D. (2015). *Change the Story, Change the Future*. Oakland, CA: Berrett-Koehler Publishers Inc.

Kruesi, M. J., Hibbs, E. D., Zahn, T. P., Keysor, C. S., Hamburger, S. D., Bartko, J. J., and Rapoport, J. L. (1992). 'A 2-year prospective follow-up study of children and adolescents with disruptive behaviour disorders. Prediction by cerebrospinal fluid 5-hydroxyindoleacetic acid, homovanillic acid, and autonomic measures?' *Archives of General Psychiatry 49*(6): 429–435.

Laing, R. D. (1959/1990). *The Divided Self*. London: Penguin.

Lancy, D. F. (2015). *The Anthropology of Childhood: Cherubs, Chattel, Changelings*, 2nd edn. New York: Cambridge University Press.

Lee, R. B. and Daly, R. (eds) (2005). *The Cambridge Encyclopedia of Hunters and Gatherers*. New York: Cambridge University Press.

LeVine, R. A., Dixon, S., LeVine, S. E., Richman, A., Keefer, C., Liederman, P. H., and Brazelton, T. B. (2008). 'The comparative study of parenting'. In R. A. LeVine and R. S.

New (eds), *Anthropology and child development: A cross-cultural reader* (pp. 55–65). New York: Wiley.

Litowitz, B. E. (1998). 'An expanded developmental line for negation: rejection, refusal, denial'. *Journal of the American Psychoanalytic Association 46*(1): 121–148.

Loy, D. R. (2002). *A Buddhist History of the West: Studies in Lack*. New York: SUNY.

Lupien, S. J., McEwen, B. S., Gunnar, M. R., and Heim, C. (2009). 'Effects of stress throughout the lifespan on the brain, behaviour and cognition'. *Nature Reviews Neuroscience 10*(6): 434–445. doi:10.1038/nrn2639

MacLean, P. D. (1990). *The Triune Brain in Evolution: Role in Paleocerebral Functions*. New York: Plenum.

Margulis, L. (1998). *Symbiotic Planet: A New Look at Evolution*. Amherst, MA: Sciencewriters.

Marshall, S. P. (1995). *Schemas in Problem Solving*. Cambridge: Cambridge University Press.

McLaughlin, K. A., Sheridan, M. A., and Nelson, C. A. (2017). 'Neglect as a violation of species-expectant experience: Neurodevelopmental consequences'. *Biological Psychiatry 82*(70): 462–471. (p. 657)

Merchant, C. (2003). *Reinventing Eden: The Fate of Nature in Western Culture*. New York: Routledge.

Millennium Ecosystem Assessment (2005). *Ecosystems and Human Well-Being: Synthesis*. Washington, DC: Island Press.

Mischel, W. and Shoda, Y. (1995). 'A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure'. *Psychological Review 102*(2): 246–268.

Narvaez, D. (2011). 'The ethics of neurobiological narratives'. *Poetics Today 32*(1): 81–106.

Narvaez, D. (2013). 'The 99%—development and socialization within an evolutionary context: Growing up to become "a good and useful human being" '. In D. Fry (ed.), *War, Peace and Human Nature: The Convergence of Evolutionary and Cultural Views* (pp. 643–672). New York: Oxford University Press.

Narvaez, D. (2014). *Neurobiology and the Development of Human Morality: Evolution, Culture and Wisdom*. New York: W. W. Norton.

Narvaez, D. (2015). 'The co-construction of virtue: Epigenetics, neurobiology and development'. In N. E. Snow (ed.), *Cultivating Virtue* (pp. 251–277). New York: Oxford University Press.

Narvaez, D. (2016). *Embodied Morality: Protectionism, Engagement and Imagination*. New York: Palgrave Macmillan.

Narvaez, D. (2018). 'Ethogenesis: Evolution, early experience and moral becoming'. In J. Graham and K. Gray (eds), *The Atlas of Moral Psychology* (pp. 451–464). New York: Guilford Press.

Narvaez, D., Bock, T., Endicott, L., and Lies, J. (2004). 'Minnesota's Community Voices and Character Education Project'. *Journal of Research in Character Education 2*: 89–112.

Narvaez, D., Braungart-Rieker, J., Miller, L., Gettler, L., and Hastings, P. (eds) (2016). *Contexts for Young Child Flourishing: Evolution, Family and Society*. New York: Oxford University Press.

Narvaez, D., Gleason, T., Wang, L., Brooks, J., Lefever, J., Cheng, A., and Centers for the Prevention of Child Neglect (2013). 'The evolved development niche: Longitudinal effects of caregiving practices on early childhood psychosocial development'. *Early Childhood Research Quarterly 28*(4): 759–773. doi:10.1016/j.ecresq.2013.07.003

Narvaez, D., Panksepp, J., Schore, A., and Gleason, T. (eds) (2013). *Evolution, Early Experience and Human Development: From Research to Practice and Policy*. New York: Oxford University Press.

Narvaez, D., Valentino, K., Fuentes, A., McKenna, J., and Gray, P. (eds) (2014). *Ancestral Landscapes in Human Evolution: Culture, Childrearing and Social Wellbeing*. New York: Oxford University Press.

Narvaez, D., Wang, L., Gleason, T., Cheng, A., Lefever, J., and Deng, L. (2013). 'The evolved developmental niche and sociomoral outcomes in Chinese three-year-olds'. *European Journal of Developmental Psychology 10*(2): 106–127.

National Research Council (2013). *U.S. Health in International Perspective: Shorter Lives, Poorer Health*. Washington, DC: National Academies Press.

Nisbett, R. and Cohen, D. (1996). *Culture of Honor*. New York: Westview Press.

Overton, W. F. (2013). 'A new paradigm for developmental science: Relationism and relational-developmental-systems'. *Applied Developmental Science 17*(2): 94–107.

Overton, W. F. (2015). 'Process and relational-developmental-systems'. In W. F. Overton and P. C. M. Molenaar (eds), *Theory and Method*, vol. 1 of *Handbook of Child Psychology and Developmental Science*, 7th edn (pp. 9–62). Hoboken, NJ: Wiley. (p. 658)

Oyama, S. (1985). *The Ontogeny of Information: Developmental Systems and Evolution*. Cambridge: Cambridge University Press.

Panksepp, J. (1998). *Affective Neuroscience: The Foundations of Human and Animal Emotions*. New York: Oxford University Press.

Panksepp, J. and Biven, L. (2011). *The Archaeology of Mind: Neuroevolutionary Origins of Human Emotions*. New York: W. W. Norton.

Paracer, S. and Ahmadjian, V. (2000). *Symbiosis*, 2nd edn. New York: Oxford University Press.

Pauly, D. (1995). 'Anecdotes and the shifting baseline syndrome of fisheries'. *Trends in Ecology and Evolution 10*(10): 430.

Perry, B. D. (2000). 'Traumatized children: How childhood trauma influences brain development'. *Child Trauma Academy*, **http://www.aaets.org/article196.htm**.

Piaget, J. (1954). *The Construction of Reality in the Child*. New York: Basic Books.

Piaget, J. (1970). *Genetic Epistemology*, trans. E. Duckworth. New York: Columbia University Press.

Piaget, J. and Inhelder, B. (1969). *The Psychology of the Child*, trans. H. Weaver. New York: Basic Books.

Pinker, S. (2011). *The Better Angels of our Nature*. New York, NY: Viking.

Pollak, S. D. and Perry, B. (2005). 'Early neglect can hinder child's relationships'. *Proceedings of the National Academy of Sciences*, Nov. 21–25.

Porges, S. W. (2011). *The Polyvagal Theory: Neurophysiologial Foundations of Emotions, Attachment, Communication, Self-Regulation*. New York: W. W. Norton.

Ramon y Cajal, S. (1909–11). *Histolgie du systeme nerveux de l'homme et des vertebres*. Paris: A. Maloine.

Rest, J. R., Narvaez, D., Bebeau, M., and Thoma, S. (1999). *Postconventional Moral Thinking: A Neo-Kohlbergian Approach*. Mahwah, NJ: Erlbaum.

Rosenblatt, A., Greenberg, J., Solomon, S., Pyszczynski, T., and Lyon, D. (1989). 'Evidence for terror management theory: I. The effects of mortality salience on reactions to those who violate or uphold cultural values'. *Journal of Personality and Social Psychology 57*(4): 681–690.

Rumelhart, D. E. (1980). 'Schemata: the building blocks of cognition'. In R. J. Spiro, B. C. Bruce, and W. F. Brewer (eds), *Theoretical Issues in Reading Comprehension* (pp. 33–58). Hillsdale, NJ: Erlbaum.

Rumelhart, D. E. and McClelland, J. L. (eds) (1986). *Parallel Distributed Processing*, 2 vols. Cambridge, MA: MIT Press.

Sandler, J. (1960). 'The background of safety'. *The International Journal of Psychoanalysis 41*: 352–356.

Schore, A. N. (2003a). *Affect Dysregulation & Disorders of the Self*. New York: W. W. Norton.

Schore, A. N. (2003b). *Affect Regulation and the Repair of the Self*. New York: W. W. Norton.

Schore, A. N. (2017). 'All our sons: The developmental neurobiology and neuroendocrinology of boys at risk'. *Infant Mental Health Journal 38*(1): 15–52. doi: 10.1002/imhj.21616

Scott, E. and Panksepp, J. (2003). 'Rough-and-tumble play in human children'. *Aggressive Behavior 29*: 539–551.

Shaver, P. R. and Mikulincer, M. (2007). 'Adult attachment strategies and the regulation of emotion'. In J. J. Gross (ed.), *Handbook of Emotion Regulation* (pp. 446–465). New York: Guilford Press.

Shepard, P. (1998). *Coming Home to the Pleistocene*, ed. F. R. Shepard. Washington, DC: Island Press/Shearwater Books.   (p. 659)

Siegel, D. (1999). *The Developing Mind*: *How Relationships and the Brain Interact to Shape Who We Are*. New York: Guilford Press.

Spinka, M., Newberry, R. C., and Bekoff, M. (2001). 'Mammalian play: training for the unexpected'. *Quarterly Review of Biology 76*: 141–168.

Spitz, R. A. (1945). 'Hospitalism; an inquiry into the genesis of psychiatric conditions in early childhood'. *Psychoanalytic Study of the Child 1*: 53–74.

Staub, E. (1992). *The Roots of Evil*. Cambridge: Cambridge University Press.

Stern, D. N. (1985). *The Interpersonal World of the Infant*. New York: Basic Books.

Stern, D. N. (2010). *Forms of Vitality: Exploring Dynamic Experience in Psychology, the Arts, Psychotherapy, and Development*. New York: Oxford University Press.

Taylor, S. E. and Crocker, J. (1981). 'Schematic bases of social information processing'. In E. T. Higgins, C. P. Herman, and M. P. Zanna (eds), *Social cognition: The Ontario symposium*, vol. 1 (pp. 89–134). Hillsdale, NJ: Erlbaum.

Trevarthen, C. (1999). 'Musicality and the Intrinsic Motive Pulse: Evidence from human psychobiology and infant communication'. In 'Rhythms, musical narrative, and the origins of human communication', *Musicae Scientiae, Special Issue, 1999–2000* (pp. 157–213). Liège: European Society for the Cognitive Sciences of Music.

Trevarthen, C. (2005). 'Stepping away from the mirror: Pride and shame in adventures of companionship—Reflections on the nature and emotional needs of infant

intersubjectivity'. In C. S. Carter, et al. (eds), *Attachment and Bonding: A New Synthesis* (pp. 55–84). Cambridge, MA: MIT Press.

Trevathan, W. R. (2011). *Human Birth: An Evolutionary Perspective*, 2nd edn. New York: Aldine de Gruyter.

Van den Berg, C. L., Hol, T., van Ree, J. M., Spruijt, B. M., Everts, H., and Koolhaas, J. M. (1999). 'Play is indispensable for an adequate development of coping with social challenges in the rats'. *Developmental Psychobiology 34*: 129–138.

Weaver, I. C., Szyf, M., and Meaney, M. J. (2002). 'From maternal care to gene expression: DNA methylation and the maternal programming of stress responses'. *Endocrine Research 28*: 699.

Westen, D. (1990). 'Towards a revised theory of borderline object relations: contributions of empirical research'. *The International Journal of Psychoanalysis 71*: 661–694.

Westen, D. and Gabbard, G. O. (2002a). 'Developments in cognitive neuroscience: I. Conflict, compromise, and connectionism'. *Journal of the American Psychoanalytic Association 50*: 53–98. doi:10.1177/00030651020500011501

Westen, D. and Gabbard, G. O. (2002b). 'Developments in cognitive neuroscience: II. Implications for theories of transference'. *Journal of the American Psychoanalytic Association 50*: 99–134. doi:10.1177/00030651020500011601

Winnicott, D. W. (1957). *The Child and the Family*. London: Tavistock.

Winnicott, D. W. (1971). *Playing and Reality*. London: Tavistock.

Winnicott, D. W. (1975). *Through Paediatrics to Psychoanalysis*. London: Hogarth Press and the Institute of Psycho-Analysis.

Woodburn, J. (1982). 'Egalitarian societies'. *Man 17*: 431–451.

Worster, D. (1994). *Nature's Economy: A History of Ecological Ideas*, 2nd edn. Cambridge: Cambridge University Press.

Notes:

(1) Many thanks to the editors for their suggestions and assistance in finalizing the chapter.

(2) The particular culture to which I refer is Western capitalism, which considers the other-than-human resources to be exploitable for human ends. Korten (2015) pointed out how the dominant culture is guided by the Sacred Money and Markets story, which contrasts with the long sustainable Sacred Life and Living Earth story that guided most societies previously and sustainable societies currently.

([3]) And accumulating evidence shows that boys are much more affected by early experience—more vulnerable over a longer period to stressors in the social environment—because of their slower maturation and less innate resilience in comparison to girls (see Schore 2017).

([4]) I discuss elsewhere how the assumptions of a selfish, aggressive human nature emerge from both species-atypical early nests, which influence world view, and the cultural narratives that have emerged to justify that misdevelopment (Narvaez 2013, 2014, 2015, 2016).

([5]) Among small-band hunter-gatherers (nomadic, immediate-return societies with few possessions), broader social practices include radical individual autonomy and radical connection to and understanding of the natural environs (e.g. Ingold 2005).

([6]) Prehistorically, and among small-band hunter-gatherers and earth-centred communities, this includes other-than-humans.

([7]) In our ancestral environments, mothering was a pleasurable experience, with mutual benefits, that necessarily was supported by the community (Hrdy 2009).

([8]) Note that womb experiences also influence brain development (e.g. Davis, Glynn, Schetter, Hobel, Chicz-Demet, and Sandman 2007) and we can inherit some traits due to the experiences our parents and grandparents had, such as anxiety (Bowers and Yehuda 2016).

([9]) I understand 'object relations' here in a broad sense, as identified by Greenberg and Mitchell (1983) and Westen (1990), without commitment to the specific claims of Klein, Fairbairn, and other original object-relations theorists.

([10]) Functional adaptation refers to the accommodation made by an individual within her lifetime. This is to be distinguished from adaptation in the evolutionary sense. The latter requires outcompeting one's rivals over multiple generations, which can only be assessed retrospectively.

([11]) There are many forms of self-regulation, such as *continent* self-control (avoiding the chocolate cake while still desiring it), versus *virtuous* self-control (not desiring the chocolate cake at all because it is not a healthy person's choice), *compliant* self-control in the face of another's aggression, and so on. Here, I focus on some basic neurobiological self-regulatory systems shaped in early life.

([12]) For reviews, see Fry 2006; Ingold 2005; Lee and Daly 2005; Narvaez 2013; Woodburn 1982.

([13]) Pinker (2011) lumps them together with more complex societies that emerged in the last 1 per cent of human genus history and calls them all more violent than present-day peoples! (Only physical violence is gauged, not institutional, emotional, or psychological

violence which are rampant in the modern world.) See data-intensive refutations of this warrior-history view in Fry 2013.

(^{14}) [Eds: A link may be drawn here with Backström's conception of love as openness (Backström, this volume).]

(^{15}) Of course, downshifting to the primitive survival systems can occur for most anyone when the stress response overwhelms other capacities.

Darcia F. Narvaez

Department of Psychology, University of Notre Dame, USA

Oxford Handbooks Online

Abstract and Keywords

This introduction provides an overview of the four chapters in this section, which explores the link between psychoanalysis and social and political theory. Each chapter examines or advocates radical change; the first appeals to psychoanalytic ideas to support radicalism concerning the basis of war and pacifism; the second deals with the organization of education; the third argues for change in psychoanalytic theory and practice by emphasizing socially radical ideas on gender; and the fourth traces the origins of radical thinking in psychoanalysis to 'Jewish modernity'. Also discussed are Sigmund Freud's Civilization, which addresses the nature of social oppression and its internalization in the superego; how society imposes normative social categories in the formation of human individuals; the role of anxiety and of defences against anxiety; contemporary populism in relation to the role of a leader and changing social expressions of inegalitarianism; and how best to contain human destructiveness.

Keywords: psychoanalysis, radicalism, education, gender, social oppression, super-ego, anxiety, populism, inegalitarianism, human destructiveness

Michael Lacewing and Richard g. t. Gipps

(p. 663) RADICALISM, in the context of social and political theory and practice, is the view that something at the very 'root' of our social and political order and institutions is wrong and needs to change. Either the understanding of human beings, their needs and possibilities, that is implicitly or explicitly assumed by social structures is mistaken or the structures fail to contribute to human flourishing or, in most cases of radical thought, both. While psychoanalysis has primarily been concerned with the individual and their mental health, a radicalism of this kind finds a clear expression in the late works of Freud, mostly famously in *Civilization and Its Discontents* (1930).[1] It is from this radicalism that discussions of the relation between psychoanalysis and social and political theory tend to begin. The four chapters included here are no exception, each exploring or

advocating radical change, either by appealing to psychoanalytic ideas to support radicalism concerning the basis of war and pacifism (Butler) and the organization of education (Rustin), or appealing to socially radical ideas on gender (themselves fed by psychoanalysis) to argue for change in psychoanalytic theory and practice (Gyler).

In his review of radical thinking in psychoanalysis, Frosh identifies its origins in 'Jewish modernity', 'a mixture of universal striving, alertness to social injustice, self-reliance, and a capacity for resilient truth-seeking that was relatively free from any impulse towards justification of existing social norms'. After Freud, the radical potential of psychoanalysis was most developed within the Berlin Psychoanalytic Society and the Frankfurt School (see Jay, this volume), the leading thinkers of which were—almost without exception—Jewish. The mid-nineteenth-century breakdown of traditional Jewish society together with the rise of anti-Semitism enabled Jewish thinkers to develop a more critical mode of thought, and Freud counted himself amongst them.

(p. 664) Psychoanalysis offers an understanding of human subjectivity that places emotions and unconscious drives centre stage, an understanding that is typically missing, or at least under-theorized, in social and political thought. For example, Butler notes that political philosophy presumes that we are disposed to live together in society, submitting to common government whether through social ties or prudential self-interest. But Freud's reflections on the death drive call this universal assumption into question. If we are also unconsciously motivated to destroy the bonds that make society possible, motivations that are acted upon most fully in war, then our understanding of how societies sustain themselves and why they break down will need reconsideration. A central recognition of psychoanalytically informed social thought is that there are contradictions in the idea of 'civilization'. What is necessary for society to be possible, e.g. sexual repression or the redirection of our destructiveness, also sets up psychological and social conditions that endanger society itself, or at least its aim of enabling positive human development. But these conditions can be neither understood nor averted by conscious rational reflection alone—hence the space for a psychoanalytic contribution to social and political theory.

Psychoanalytic ideas thus provide a lens for reading contemporary social and political trends and events in terms of the presence and force of such unconscious emotions and drives, enabling us to better know ourselves. A brief selection of these ideas from the four chapters is collected here:

1. The primary focus of Freud's *Civilization* was the nature of social oppression and its internalization in the superego. With the resulting sexual repression no longer quite what it used to be, the vicissitudes of the relation of society to pleasure, and in particular the idea of the 'management' of pleasure, were explored by the Frankfurt School and are revisited here by Frosh. The question remains, how should we understand and respond to society's encouragement of (demand for?) the search for pleasure?

2. In the formation of human individuals, society may be said to impose normative social categories. Perhaps none is more deep-rooted than that of gender. Gyler argues that, despite the social transformations of the last fifty years, we continue to understand and elaborate our subjective experience in terms of normative gender categories. What response does this require from us at this stage? While her focus is on the response of psychoanalysts themselves, the question generalizes.

3. Central to the psychodynamic understanding of the mind is the role of anxiety and of defences against anxiety. These defences can be social as well as individual in form. Rustin argues that in addition to anxieties and defences concerning love and hate, deriving from the life and death drives, psychoanalytic theory needs to find a central place for anxieties and defences deriving from a distinct epistemophilic drive. The concept of epistemic anxiety has important implications for the organization of education, as evidenced by recent social events, including the 2016 debates over Brexit in the UK and the election campaign of Donald Trump, in which social antagonisms around educational advantage played an important role.

(p. 665) **4.** We may relate aspects of contemporary populism to a question raised by Butler (Chapter 42) concerning the role of a leader and changing social expressions of inegalitarianism: 'If we say that an elected official has licensed a new wave of misogyny, or that he has made widespread racism permissible, what sort of agency do we attribute to him? Was it there all along, or has he brought it into being? Or was it there in certain forms and now his speech and action give it new forms?'

5. Among the defences against anxiety and our natural tendency to envy (Klein) are various phantasies and projections concerning our relation to authority and to the social 'other' (race, class, immigrant . . .). What kinds of social arrangements and institutions could ameliorate anxiety and envy, to provide some form of containment (Bion) for them? Might such institutions be opposed because they remind us of the vulnerability that gives rise to anxiety in the first place?

6. Finally, how should we understand human destructiveness and how best to contain it?

While psychoanalysis raises these questions, a further theme of these chapters is whether psychoanalysis has sufficient resources to answer them. Does its radicalism run out at this point? While the critical resources of psychoanalysis are apparent, its potential to contribute positively to radical social and political theory is less obvious. First, in its own clinical practice and theory, it has often been conservative. Gyler makes the case forcefully concerning gender identity and inequality. While contemporary psychoanalytic thought has turned away from a focus on gender, it nevertheless retains a traditional association of masculinity with reason and culture and femininity with emotion and nature, or again, the feminine continues to be conceived as 'lack' or 'absence'. There is little positive contribution to a transformation of this patriarchal symbolic order, demonstrated for example by how little psychoanalysis has had to say about transgender.

Frosh (Chapter 39), drawing on previous work by Rustin, observes that psychoanalysis can enable us to think how 'specific modes of social organization can moderate or produce states of increased vulnerability or greater security, exaggerated violence or

more resilient possibilities of nonviolent, reparative practice'. Butler (Chapter 42) similarly comments that 'for societies that wish to organize themselves on the basis of a principled opposition to cruelty, that is, to war or to the death penalty, there first has to be a mindfulness of the destructive potential that inheres in all group formations'. But, and this is a second reservation concerning the radical contribution of psychoanalysis, these responses contribute towards social improvement, rather than pointing us towards anything more revolutionary.

Third, one clear line of thought in Freud about how to respond to the antisocial tendencies of human beings—our anxiety, envy, and destructiveness—is to harness and strengthen the bonds of life and love, the forces of Eros, that hold us together. But even Freud was doubtful whether this can suffice. For example, in his exchange with Einstein, 'Why War?', discussed by Frosh and Butler, he agrees with Einstein's emphasis on the need for an international agency that can limit state sovereignty to prevent the outbreak of war. But for such an agency to work and contain aggression, it would need to rest on (p. 666) widespread sentiments favouring internationalism over nationalism. For this, a 'cultural evolution' is needed—but is it possible?

Butler concludes her chapter by identifying the importance of critical thought, which among other things is 'a form of judgement, that knows how and when to check that destructiveness and that has the power to do so'. Critical thought comes under attack in certain forms of 'mass psychology', but she argues, '[s]uch judgement is crucial to the operation of politics, and one task of political theory is to establish the conditions of its possible exercise'. If our aggression is inescapable, then perhaps it can be rallied to the cause of life through its expression in critical thought—we rally our hatred against war and destruction to adopt an aggressive form of pacifism. With these last thoughts, a radical political response finds some place in psychoanalytic thought.

Reference

Freud, S. (1930). 'Civilization and its discontents'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 21 (pp. 57–146). London: Hogarth Press.

Notes:

(1) Radicalism, in addition, has frequently advocated revolutionary activity to bring about the necessary change. This is perhaps less prominent in psychoanalytic thought, though it finds expression in some of the works of the Frankfurt School and of Žižek.

> **Michael Lacewing**
>
> Dr Michael Lacewing is a former Vice-Principal Academic and Reader in Philosophy at Heythrop College, London, an Honorary Reader in Research Department of Clinical, Educational and Health Psychology at University College, London, and a teacher of philosophy and theology at Christ's Hospital School, Sussex. He has

published widely in philosophy of psychoanalysis, metaethics and moral psychology, alongside writing textbooks for A level philosophy and training in Philosophy for Children (P4C).

Richard G.T. Gipps

Dr Richard Gipps is a clinical psychologist in private psychotherapy practice in Oxford, UK, and an associate of the Philosophy Faculty at the University of Oxford. He convenes the Philosophy Special Interest Group of the Institute of Psychoanalysis, the Oxford Interdisciplinary Seminars in Psychoanalysis, and the Making the Unconscious Conscious seminar series. His research interests lie in psychoanalysis, psychosis, existential phenomenology, and Wittgenstein.

Oxford Handbooks Online

Abstract and Keywords

This chapter is concerned with the contribution that psychoanalysis has made to progressive political thought. It argues that despite, alongside, or in tension with the more conservative, psychologically 'reductive' side of psychoanalytic politics, there is a very challenging radical strand. On the whole, once the Berlin Institute of Psychoanalysis was destroyed by Nazism, it found its strongholds outside the main psychoanalytic movement, for example in the works of philosophers and social theorists from Herbert Marcuse to Judith Butler; and this is one of the issues that needs to be addressed as part of the question of whether this radicalism is truly 'psychoanalytic'. Starting with Freud, and taking seriously the contribution of social theorists influenced by Klein and Lacan, the chapter suggests that psychoanalysis offers a vocabulary for, and orientation towards, subjectivity that is not otherwise highly developed in political thought.

Keywords: conservatism, radical politics, social theory, subjectivity, Freud, Klein, Lacan

Introduction

It can be claimed that psychoanalysis has always operated in a political domain. From his earliest formulations onwards, Freud was interested in how individuals might find themselves at odds with their society, specifically through the opposition between sexual drives and social repression. The question of human freedom in relation to a fundamentally constraining social world (the world of the 'reality principle') has recurred throughout the history of psychoanalysis. The political engagement implicit in this ostensibly libertarian query—how much latitude can the individual be allowed in a

regulated society?—is augmented by a number of other strands, ranging from the conservatism visible in the privatization of clinical practice, through social welfarism and radical, socialist, or Marxist activism and their contemporary post-Marxist forms.

Although most of these tendencies have competed with one another throughout the history of psychoanalysis, a loosely chronological tracing of these different political directions is possible. Freud's 'social awareness' and his commitment to social democratic practice was manifested in his speech to the 1918 International Psychoanalytic Congress in Budapest advocating free treatment for 'the wider social strata' (Freud 1919: 166), which led to the founding of free psychoanalytic clinics in Berlin, Vienna, and elsewhere in the 1920s and 1930s (Danto 2005). The promise and limitations of the psychoanalytic radicalism of the time, which was wrecked by the onset of Nazism and the collapse of German psychoanalysis (Frosh 2005), gave way after the Second World War to the more normative practices of ego-psychological and object-relations work in the USA and Britain. From the 1960s onwards, however, there has been a return to various politically active strands in psychoanalytic thought, such as in certain uses of Lacan (Stavrakakis 2007); in the profound challenge to psychoanalysis that came from feminism (Goldner 2003); and in more recent engagements with psychoanalysis by queer and postcolonial theorists (Giffney and Watson 2017; Khanna 2004). On the other hand, psychoanalysis as a practice remains quite conservative and at times (for instance in Latin America during the dictatorships of the late twentieth century) there has been collusion between psychoanalytic institutions and oppressive social regimes, oriented around a cult of 'neutrality' and a familial ideology that was easily appropriated by authoritarian rulers (Rubin et al. 2016).

The conservative elements in psychoanalysis are genuine and would benefit from extended treatment on their own. They have roots in a variety of sources, including Freud's personal attitudes (especially towards women and Bolshevism—e.g. Roudinesco 2016; Makari 2008); the strong, yet relatively unacknowledged implication of much early psychoanalytic thinking in colonial assumptions and racializations, as reflected in Freud's *Totem and Taboo* but also in general psychoanalytic notions of 'primitivity' (Brickman 2003; Freud 1913; Frosh 2017); the medicalization of psychoanalysis under the influence of Ernest Jones and American psychoanalysis (itself driven by the search for professional respectability and fear of 'quackery'—see Makari 2008; Zaretsky 2015); the increasingly bourgeois nature of psychoanalysis as it settled down, especially post-Second World War, into a middle-class profession never fully integrated into the National Health Service in Britain or public health provision elsewhere (Ryan 2017); and a conceptual affiliation to psychology or psychologism, with its characteristic 'reduction' of complex social experiences to 'internal' psychological events (Frosh 1989). In this regard, the emergence of psychoanalysis in the late nineteenth century at the same time as psychology and sociology meant that it participated in the epistemological divisions between what was legitimately the arena of the sociological—the operation of social forces at a structural level—and what was the domain of the psychological—individual behaviour and the 'inner' attributes that lie behind it. Psychoanalysis generally affiliated itself to the psychological domain (this is indicated even by its name, which has never been, as some

might have liked, 'socioanalysis') and contested the grounds of theory and practice with academic psychology and clinical psychology as well as with psychiatry, with their associated individual-centred treatments such as behaviour and cognitive behaviour therapy and, of course, psychopharmacology. For psychology, the medium of engagement with the social and political was largely 'social psychology', understood primarily as the investigation of how social context affects individual psychology (attitudes, group processes, etc.). Similarly, 'official' psychoanalysis largely limited its forays into the sociopolitical world to accounts of the requirements for 'adaptation' (Jacoby 1975) and investigations of the impact of trauma, especially post-Holocaust and then in relation to sexual abuse once it emerged as a major issue in the 1980s. Despite paying homage to Freud's great late social texts (The Future of an Illusion (1927), Civilization and its Discontents (1930), Moses and Monotheism (1939)), psychoanalysis has largely focused on the struggles of individuals to survive their tumultuous inner world and often difficult early circumstances. This is not unwarranted given the therapeutic focus of its practice, and has produced a rich array of concepts relating to what might be required of a caring society in order for people to develop the internal resources and capacities allowing them to manage themselves (Winnicott's (1965) notions of what a baby might need from the maternal 'environment' are particularly expressive and influential here); but it has had relatively little to offer in relation to political thought.

If the tendency towards a more 'conforming' or at least individualizing understanding of the relationship between people and their society has been characteristic of mainstream psychoanalytic thinking in the post-war period, there was plenty of radicalism in its early years, expressed most influentially in the Berlin Psychoanalytic Institute of the 1920s and very early 1930s (Frosh 2005; Goggin and Goggin 2001; Jacoby 1983). The Berlin Institute was a remarkable hotbed of psychoanalytic creativity that also housed a large group of socialist intellectuals; and its famous 'Children's Seminar', run mainly by Otto Fenichel, was an explicit attempt to reconcile the 'dialectical' psychological theory of Freud (the dialectic being produced by the conflicting psychic forces either of sexuality and 'ego preservative' drives, as in Freud's first theory, or life drives and death drives, as in his post-First World War reformulation of drive theory) and Marxist social-revolutionary dialectics. This left a strong legacy, despite the corruption and destruction of the Berlin Institute by Nazism and the exile of its adherents. It has been a place to look back to and locate a genealogy of psychoanalytic radicalism that depends less on Freud (though his social texts still form the basis of most psychoanalytic radical thought), and more on the mix of theory and social practice that the Berliners advocated. It is with this 'tradition', if that is not too strong a word, that this chapter is primarily concerned. That is, despite, alongside, or in tension with the more conservative, psychologically 'reductive' side of psychoanalytic political thought, there is a very challenging radical strand. On the whole, once the Berlin Institute went under, it found its strongholds outside the main psychoanalytic movement, for example in the works of philosophers and social theorists from Herbert Marcuse to Judith Butler; and this is one of the issues that needs to be addressed as part of the question of whether this radicalism is truly 'psychoanalytic'. Or perhaps more usefully, the question should be: what is psychoanalytic about this

Subscriber: University of Otago; date: 17 December 2018

philosophical strand of radicalism, and what can it draw from psychoanalysis that can contribute to the radical political endeavour? To jump forwards quite a lot, the simple answer to this is that psychoanalysis offers a vocabulary for, and orientation towards, subjectivity that is not otherwise highly developed in political thought; but what this means, and what its promise and limitations might be, still needs to be unravelled.

Sexual Emancipation

One way of reading Freud is as a representative of what might be called 'Jewish modernity' characterized by the explosive release of Jewish intellectuals from the constraints of traditional societies, alongside the continued enforcement of their marginality by ever-evolving anti-Semitism. Enzo Traverso (2016: 9) describes how access to citizenship—'emancipation', which occurred in Austria only in 1867, when Freud was eleven years old—began a process of dissolution of Jewish community life that was nevertheless highly constrained: 'From this turn on, the marginality of Jews was more a question of the attitude of the world around them than of their own desire to preserve a separate life'. Modern European anti-Semitism imposed from the outside what Jewish community self-regulation, albeit aided by Christian anti-Semitism, had previously preserved from the inside. Traverso comments (2016: 9), 'This is the source of the mixture of particularism and cosmopolitanism that characterizes Jewish modernity'. It also characterizes Freud, who as a highly educated professional of the late nineteenth century might have expected to be entitled to full acceptance within society, yet who as a Jew experienced exclusion on a regular and thorough basis. The consequence, according to Freud himself as well as in line with the broad thrust of Jewish modernity, was a mixture of universal striving, alertness to social injustice, self-reliance, and a capacity for resilient truth-seeking that was relatively free from any impulse towards justification of existing social norms. That is to say, Jewish modernity bred a *critical* impulse. Looking back to the 1890s on his seventieth birthday in 1926, Freud (1961: 368) famously wrote, 'Because I was a Jew, I found myself free of many prejudices which restrict others in the use of the intellect; as a Jew I was prepared to be in the opposition and to renounce agreement with the "compact majority"'. The impulse here is towards independence of mind fuelled by a position of marginality that is uncomfortable, for sure, but also has the advantage of leaving open the space for critical thought.

Freud's (1908) paper, '"Civilized" Sexual Morality and Modern Nervous Illness', with its argument that neurosis is in large part caused by the hypocritical relations governing sexuality in the Europe of the early twentieth century, can be understood as just such a critical intervention in the social and political mores of his time. His argument here is potentially quite seditious: sexual repression might be necessary for the maintenance of the social order ('civilization'), but the 'sacrifices' it imposes on people are such as to cause them suffering pand thereby endanger society, or at least its 'cultural aim' (1908: 180). The various restrictions of sexuality, for instance in sexual abstinence and

monogamy, which supposedly moderate the passions that could undermine social relations, are themselves damaging of society because they cause neurosis. Freud acknowledges the possibility that the 'damage done' in this way might be necessary for the well-being of society as a whole, but he is not convinced: 'I must confess that I am unable to balance gain against loss correctly on this point, but I could advance a great many more considerations on the side of the loss' (1908: 195). This means not only that suffering is imposed on individuals, but society as a whole is damaged by the repression of sexuality, with the implication that radical social change might be required in order to achieve a reduction in neurosis and also the creation of a healthy social environment.

If '"Civilized" Sexual Morality and Modern Nervous Illness' marks one beginning of 'Political Freud' (Zaretsky 2015), *Civilization and its Discontents* (Freud 1930) is the high point of his explicit engagement with social and sociological issues. In this book, written as Europe was beginning to smell the gathering vapours of destruction again, Freud returns to the issue of whether the suffering of individuals is a necessary consequence of their position as social subjects. This time, he was armed with his new concept of the death drive, which Butler (this volume) has examined closely and shown to have constructive as well as destructive prospects. As in the earlier piece on sexual morality, 'civilization' is presented as a necessary evil. Without it, people would be subjected to various uncontrollable forces, not least those that come from their own impulses— impulses that from prehistoric times have always been constituted as much by violence as by love. *With* civilization, however, comes privation and constraints on freedom that are felt to be unjust: 'The liberty of the individual', writes Freud (1930: 94), 'is no gift of civilization'. The political consequence of this is that there is always the potential for revolt against society. Sometimes this is because of the actual injustice of the social order, so the push for freedom may 'prove favourable to a further development of civilization'— an indication that Freud was not in principle averse to the possibility of political rebellion. But sometimes what is being expressed in this way is a general 'hostility to civilization', a reaction against the very idea that there should be any control of the drives at all. In this cautiously expressed formula, Freud attests to a kind of liberal concern for moderation. On one side, revolt might be atavistic and wish-fulfilling, producing the irrationalities of destructiveness that negate all cultural gains. On the other side, there are times when it is all too much and it is understandable and even socially constructive that people might seek to throw off some of their chains.

Balancing optimistic and pessimistic prospects for political resistance is characteristic of Freud's thought, and is enunciated in the final lines of *Civilization and its Discontents* as a never-ending battle between the death drive and Eros. It can be seen working its way right through that text, particularly in relation to the formation of the superego as an internalization of violence, strikingly imagined by Freud as 'like a garrison in a conquered city' (1930: 124). Here the superego is presented along a specifically political, or rather *military* model, with the 'garrison' being the mechanism of colonization that maintains order on behalf of a hostile external entity. The 'pessimistic' side of this is that the internalization of the external order in this way means that the 'law' that is the abstract version of the father in the Oedipus complex (Laplanche and Pontalis (1967: 286) give as

one rendering of the Oedipus complex 'the triangular structure constituted by the child, the child's natural object and the bearer of the law') comes to regulate the subject from the 'inside', thus freeing society from the requirement to become so authoritarian that it exhausts itself in repressive activities. Although he did not use this language, Freud's account here prefigures the idea of the 'administered society' that emerges in later critical theory (Adorno 1967): resistance to social oppression is made all the harder when social norms are experienced as natural because they are thoroughly internalized, and the subject becomes self-regulating rather than simply subjugated to social norms. With the formation of the superego, the subject always feels guilty and at fault: as Freud notes (1930: 128): 'A threatened external unhappiness—loss of love and punishment on the part of the external authority—has been exchanged for a permanent internal unhappiness, for the tension of the sense of guilt'. If there is a more 'optimistic' element that can be read into this material, it is in the way that the internalization of social norms can also humanize them, meaning both that collective identifications are possible and that the violence of the social is intimately connected to its loving elements. One consequence of this, as Butler (this volume) has been exploring in her recent work (see also Butler 2015), is the possibility of solidarity amongst those who have much to lose, or indeed have already had to deal with the losses imposed by social regulation. One might add to this the psychoanalytic claim that emancipation is produced by a form of *knowledge*. Insight into the internalized norms of the social order can activate resistance to them when the life drives become aware of just how damaging these norms can be. This might even be understood to be a justification for Freud's sociological writings as a mode of 'applied' psychoanalysis: understanding of the psychoanalytic kind is a step towards becoming free.

Wriggling out of the pessimistic position and articulating a more viable emancipatory one has been one of the ambitions of much psychoanalytic radicalism, epitomized especially in the 'libertarian' stance taken early on by Wilhelm Reich and then by Herbert Marcuse. There is quite a lot that can be rescued from Reich's troubled, even tragic, trajectory through psychoanalysis to oblivion, one that has been written about from many perspectives, including through the scathing irony of Philip Rieff (1966). Rieff saw Reich as a misbegotten mystic, an advocate of the 'ecstatic' attitude that could lead only to the decline of the true psychoanalytic contribution, which for Rieff was the maintenance of a capacity to 'live with one's ailments'. In retrospect, this is accurate and yet also too harsh, and Rieff himself clearly recognized this because at the end of his portrait of Reich he reinstates him as a kind of priest of love.

> Reich yearned for a revolution of mood. When both Freudianism and Marxism failed to make that revolution, he invented his own, although it came to have an almost entirely private use … Nevertheless, though publicly labelled an eccentric, Reich was anything but a fool. On the contrary, he was wise enough to know that love is the supreme form of energy … pathetic as Reich's method was, it is well to be reminded that love is the ultimate power

(Rieff 1966: 160–1).

It is perhaps this utopian element in Reich's thought that has continued to catch people's imagination. Reich's ideas, especially from his early years, have been provocative and often influential on those who are impatient with the often formulaic pessimism of psychoanalysis about therapeutic and social change. Reich's eventually eschatological view of life and death (orgone energy as the erotic stuff of the universe) had its underpinnings in an understanding of psychic resistance in which the sexual energy of the free subject is constrained by the internalization of family structures in the form of 'character'. This means that what are structured into the social world as authoritarian, especially fascist, living conditions are also experienced in the individual's psychic life as a personality or character structure that inhibits and suppresses liveliness, becoming, in both political and psychoanalytic senses, *repressive* (Reich 1946). This not only makes the supposedly 'inner', psychological world just as important as the outer one in relation to continued oppression and resistance to it, it additionally shows up the bonds between them, their inextricability. *Sexual* revolution is as necessary as political revolution if anything is to change, because the one without the other leaves the authoritarian social structures unaltered. There is thus a profound stirring of the psychoanalytic pot going on in Reich's early work, in which the ideological conditions of domination become structured into sexual life.

Traces of this libertarian perspective can be seen in Marcuse's (1955) *Eros and Civilization*, despite Marcuse's more systematic and thoroughgoing exploration of both Freudian and Marxist theory (see Jay, this volume). Once again there is the idea that the drives are fundamental, and that—as in Freud—they are met by a certain degree of social resistance. However, whereas Freud emphasized the *necessity* of this social resistance for maintenance of an ordered society, the radical critics argued that such 'repressive' forces are in fact, under capitalism, structures of *domination* and the idea that they are anything else is part of the ideological process whereby subjects are regulated and 'administered' or controlled. In late modernity (Marcuse was writing in the mid-1950s) this system of administration is brought to the boil by the possibility of technological alleviation of the supposedly real conditions of nature; that is, there is an increased potential for libidinal release produced by technological advances. However, to counteract this the 'administered society' turns to the management of pleasure itself; which is to say, the contemporary mode of domination is less by active and overt force, and more by the subtle seduction of the subject into its processes of consumption. Advertising replaces armies (though one has to question the global accuracy of this assertion); as Slavoj Žižek (2006) never tires of reiterating, 'enjoyment' ceases to be a way of resisting social constraints and instead becomes a social injunction. We are required to enjoy, to take possession of the commodities available under capitalism and to use them as substitutes for freedom.

It is in this way—through the management of pleasure—that society is perpetuated in the interests of capital. This resonates strongly with later work on how, for example, capitalism provokes and frustrates desire at one and the same time, so that each failed

attempt at satisfaction drives the subject to consume more. What no one can bear, neither subject nor society as a whole, is the possibility of coming face to face with desire and the necessary impossibility of its fulfilment—otherwise it would putatively cease to be desire, a point discussed later. Instead, this process of administration reassures the subject that while it may be dissatisfied now, salvation is just around the corner. Žižek, as it happens, has an important and characteristically provocative response to the revelation of the administered society and the place of psychoanalysis within it. 'Traditionally', he writes (2006: 304):

> psychoanalysis was expected to allow the patient to overcome the obstacles which denied him or her access to 'normal' sexual enjoyment; today, however, when we are bombarded from all sides by different versions of the superego injunction 'Enjoy!', from direct enjoyment of sexual performance to enjoyment of professional achievement or spiritual awakening, we should move to a more radical level: today, psychoanalysis is the only discourse in which you are allowed *not* to enjoy (as opposed to 'not allowed to enjoy').

This is a new version of the old idea that psychoanalysis can help remove illusions, including ideological ones. One question, in addition to the obvious one about efficacy (does psychoanalysis really succeed in this?—a different question from the one about therapeutic utility, but entertainingly parallel to it), is what we might be left with if this were to occur. Freed of ideological mystification, what does one then see? Moreover, is it the injunction to *enjoy* that characterizes contemporary western society (and one needs to note the very particular cultural location of this claim, which seems to be presented by Žižek as if it applies to all societies, and to cross over categories of gender, ethnicity, and desire) or are there other competing injunctions? To give one example, Isin (2004) portrays the generalized contemporary subject as a subject of *anxiety*, always dissatisfied and living in fear of catastrophe. This 'neurotic citizen' is produced as such by governing practices that treat the subject 'as someone who is anxious, under stress and increasingly insecure and is asked to manage its neurosis' (2004: 225). The neurotic subject, Isin claims (2004: 225), 'is one whose anxieties and insecurities are objects of government not in order to *cure* or *eliminate* such states but to manage them'. Citizenship becomes a space for the appeasing of anxieties that have themselves been created as part of the process of governing; this promotes paranoia and a hunt for security as well as a constant process of self-monitoring that is deliberately induced to block political action (Frosh 2017). While this has moved quite a way from Marcuse's original analysis, there is a recognizable genealogy in which identification of how desires are 'administered' develops into a fuller account of the management of anxiety and pleasure in a surveillance society, with its attendant implications for control and—ultimately, if we are lucky and persistent —resistance.

War, Barbarism, and Rationality

The exchange between Freud and Albert Einstein published as *Why War?* in 1933—an ironic year, given the election of Hitler to power in Germany—is a strange piece that as Butler (this volume) also notes, compresses some productive reflections on war and peace into a compelling, dialogic space. *Why War?* presents the spectacle of these two great Jewish intellectuals speculating about human destructiveness and the reasons why rational solutions to violence do not take hold. They seem to be in agreement that there must be something equivalent to a death drive, an inbuilt tendency in the human subject to choose to move against peace. For Einstein, this takes the form of 'a lust for hatred and destruction'. He opines that, 'In normal times, this passion exists in a latent state, it emerges only in unusual circumstances; but it is a comparatively easy task to call it into play and raise it to the power of a collective psychosis' (Einstein and Freud 1933: 201). Freud seems disconcerted by this ('You have taken the wind out of my sails' (Einstein and Freud 1933: 203)) as if surprised that this man who claims to know little about psychology could so easily come up intuitively with an idea that Freud himself has struggled towards over decades: that inside each person there is an urge for destruction, upon which warmongers and aggressive nation states can build. Freud uses the space of the exchange to produce an account of the death drive that emphasizes its biological basis and also its *protective* function: the death drive is to be understood as a compulsion to return to rest (like all drives in the Freudian scheme, one might argue) that becomes externally violent as a way of protecting the organism against its tendency to self-destruct. Perhaps this overly complicated account, which belies the supposed aim of the exchange to make Freud's thinking on war widely accessible, reflects how Einstein cuts the ground from under Freud's feet while apparently setting up the discussion in such a way that Freud can lead it. Not knowing anything about psychology, or so he claims, and enthralled by the great perspicacity of the discoverer of the unconscious, Einstein nevertheless, naively, comes up with precisely the same theory that Freud has laboured his whole life to produce. *Anyone can do it* is the unconscious message here: Einstein is as good a psychologist as Freud. The problem for Freud then is that he can only 'confirm all you have said', and in so doing make his own additional contribution either excessively complex for this public and non-technical interchange, or trivial, no better than the amateur can produce without trying. The famous Jewish joke about the response to an attempt to describe Einstein's theory—'from *this* he makes a living?'—is turned by Einstein onto Freud.

It may be that what is going on here is that Freud is rebelling against the reduction of the death drive to the observation that people tend to be destructive, or even to widely held religious and biological notions about the animalistic elements in 'human nature'. For Freud, the central point about the death drive is not that there is a source of evil within each person, or even that each one of us is fundamentally antisocial (or at least *asocial*), even though there are traces of these views in his account. It is rather that there is a universal wish for the solace of dissolution, what he refers to as the 'nirvana' principle

(Freud 1920: 56), characterized by the pull of the subject towards inactivity and complete rest. Most of all, there is the structure of repetition, in which things keep coming into view and then disappearing again, like the child's cotton reel in the famous fort-da game of Freud's grandson Ernst (1920: 15). The death drive *becomes* violent and destructive though its projection outwards into the world, and *Civilization and its Discontents* is in large part a meditation on the consequences of this and on what might be done to protect humanity against these consequences. But the death drive is also a way of thinking about those things that cannot easily be wished away, but keep returning to plague us until we can find ways of dealing with them—sadness and loss, failure and trauma, war and unbearable suffering. We know through how we deal with these things that something stirs inside us to wreck everything we depend on; the big question is what can be done with this destructive urge, how can it be contested or managed? None of this is easy to theorize, nor can it be reduced to the commonsense notion of a 'lust for destruction and hate' even if this might capture some of its force.

For Freud, it is the life drive that embodies hope, 'eternal Eros' as he calls it, drawing on the vocabulary of love that sustained him over the years. The radical possibility here would be that the force of love, or more broadly the libidinal energy that elaborates life and fuels the capacity of humans to come together to create and procreate, can be harnessed in opposition to destructiveness, one force against another, the two 'immortal adversaries' locked in combat. Others, however, have focused differently on the way in which destructiveness can be thwarted, while still taking the death drive seriously. The keenest example is that of Kleinian psychoanalysis, where the inborn deathliness of the subject is taken as a given, and the trajectory of thought is towards what can be done to live with this. For Klein, the death drive is manifested psychologically in the state of envy, which 'is an oral-sadistic and anal-sadistic expression of destructive impulses, operative from the beginning of life, and … has a constitutional basis' (Klein 1957: 176). Envy is unavoidable and it makes no difference if the object of envy is generous or persecutory. If the former (as with the breast that the infant experiences as full and giving), it can be hated for the fact that it possesses such riches; and if the latter (the mean and grudging breast), it is hated because it refuses to give what it should: the infant, Klein writes (1957: 180), 'feels that the gratification of which he was deprived has been kept for itself by the breast that frustrated him'. When envy is intense, the perception of a good object can be as painful as that of a bad one, for the better it is the more it gives rise to envious wishes. Envy therefore destroys hope, and good things are poisoned by it; its regular appearance in psychotherapy is both a necessary focus for work and a profound threat to progress.

This description of envy seems to offer little possibility for radical resistance or for reconstruction of a broken world. It is also a reason why Kleinianism is seen as biologistic: everything comes from within, and whatever the nature of the external environment, it is the force of these constitutional drives that determine the structure and phenomenology of psychic life. This characterization of Kleinian theory as strongly focused on the 'inner world' has a lot of truth, and is a compelling description of much Kleinian clinical practice (Hinshelwood 1994). Nevertheless, in recognizing the ubiquity

Subscriber: University of Otago; date: 17 December 2018

of human destructiveness, Kleinian psychoanalysis opens out to something else—the issue of what kind of society might be constructed in order to ameliorate this, or at least to make it possible for people to live together peacefully. Indeed, it might be argued that it reflects one of two possible routes into the question of dealing with violence, one of them the ameliorative route and the other based on an acknowledgement of needs and of the conditions of sociality that promote the possible meeting of those needs. This differentiation, which perhaps does not have to be an absolute one, produces some slightly unexpected genealogies. On the Kleinian side, for example, there is acknowledgement of the role of the environment in mitigating envy and easing destructive impulses, even though these impulses can never be fully removed or restrained. As Klein herself notes in relation to early infancy, certain kinds of 'good' experience can begin an integrative process that allows envious urges to be offset by loving ones.

> If the undisturbed enjoyment in being fed is frequently experienced, the introjection of the good breast comes about with relative security. A full gratification at the breast means that the infant feels he has received from his loved object a unique gift which he wants to keep. This is the basis of gratitude (1957: 188).

More powerfully, the Kleinian idea of 'reparation' has been picked up to express a way of coming to terms with destructiveness that has resonance both in psychoanalysis and in social relations. Applied to individuals, reparation refers to the way in which the ego 'feels impelled (and I can now add, impelled by its identification with the good object) to make restitution for all the sadistic attacks that it has launched on that object' (Klein 1935: 149). The issue here is of how to repair what has been damaged, and specifically what each of us feels we have done violence to. It is when we attack what we also need and love that depression follows, and it is under such circumstances that reparation is called for. Reparation rebuilds the world after destruction and as such, as Michael Rustin (1995) points out in relation to Kleinian theory as a whole, it is a 'positive', integrative mechanism: it starts from fragmentation and paranoid splitting (evil versus good; envy versus gratitude) and brings together the damage done with the moral responsibility of the one who has done that damage, so constructing an impulse towards rebuilding and repair. On the other hand, there is something symptomatic in the use of the terminology of reparation, invented at a historical moment in which it had by no means a solely integrative and healing set of associations. Lyndsey Stonebridge (1998) has demonstrated that 'reparations' in its original, post-First World War context, not only meant repairing damage, but also had connotations of being unfair and punitive, and these 'violent' connotations left some trace in psychoanalysis itself. Reparation is one of those complex psychoanalytic concepts that contains many ambiguities, including a tinge, at least, of hostility alongside its dominant ameliorative connotations. As Stonebridge argues, some of this can be seen in Kleinian analyses of creative artwork (understood as produced largely through reparative processes—Segal 1990), which contain their own dynamic of

destruction and reparation, although it is perhaps more arguable whether the hostile component of 'reparations' is maintained within contemporary Kleinian clinical thinking.

The idea of reparation has had creative resonances for writers seeking to find a way of expressing the possibility of a social world constructed around conditions of care without idealization. For Rustin (1991), for example, working sociologically in the Kleinian tradition, the issue has long been how to create a model of social reform that will seriously address the fragility of the social order in the face of potential destructiveness, and will maximize the conditions for responding to the need for care. Faced with undeniable tendencies towards splitting and destructiveness, what are the requirements for social conditions that will contain these tendencies sufficiently for them to become less fearsome, less potent, and for the equally important yet often subjugated impulses towards gratitude and love to come to the fore? Perhaps the most obvious, and probably most successful, instance of this working out well is in the construction of the welfare state in Britain after the Second World War. Destruction had certainly made itself felt in acts of barbarism and uncontrolled brutality that overwhelmed much of the world— especially the supposedly 'civilized' world of Freud's imagination—to a degree perhaps never seen before. In response, rebuilding the shattered economy of the country in a context of austerity and continuing rationing, the Labour government of 1945–50 introduced profound and lasting reforms in education, welfare, and perhaps most of all in healthcare. Psychoanalytically and possibly politically too, though this is not really the language of politics, one can understand this as a way of making reparation to the citizens of the country for what they had gone through. The force of destruction had been felt as vividly as it could imaginably be; the possibility of further devastation through nuclear war seemed very real; and while this demanded political reaction at the level of confrontation of this destruction, it also needed acts that would model and mould a society built around care rather than further suffering. This is exactly what happened in the formation of the UK's National Health Service, and is also perhaps one reason amongst many that social welfare is so virulently attacked by neo-liberals and other right-wing forces: it reminds us of our dependency and vulnerability, of how much we need protection and exactly why that should be so.

The legacy of this thinking in cultures of care is important (Hollway 2006), but it is also worth noting how it resonates with other work on violence. Here as in many other places, Judith Butler's influence is significant. Butler regards violence as ubiquitous and endemic to the formation of the human subject, experienced in infancy not only through acts of commission or omission (abuse, neglect) but in a 'routine' way through the forceful imposition of social categories on the subject.

> We are all at least partially formed through violence. We are given genders or
> social categories, against our will, and these categories confer intelligibility or
> recognizability, which means that they also communicate what the social risks of
> unintelligibility or partial intelligibility might be

(Butler 2009: 167).

Butler is stressing here both that we are all impinged on in a more or less violent way, all submitted to a regime of dependency that makes us vulnerable, and that the incorporation of a human subject into sociality is a 'violent' process in which certain pregiven, 'iterable' structures (in the quotation above those of gender) have to be accepted. This significantly extends the usual psychoanalytic notion of what violence might be, and it draws out the political implications of the social regime of violence, linking Butler's thinking with that of Marcuse and the Frankfurt School. It also allows Butler to articulate the responsibility that falls on each human subject to resist reproducing the violence that we are so fully 'mired' in. 'It may be', she writes (2009: 167), 'that precisely because one is formed through violence, the responsibility not to repeat the violence of one's formation is all the more pressing and important'. We know about violence in our very formation, in the fundamental depths of our lives, just as we know about vulnerability and dependency through the earliest and most deeply rooted neediness of infantile experience. In response, we can either 'moralize' our situation and inflict violence on others, or we can 'struggle with the idea of non-violence in the midst of an encounter with social violence as well as with [our] own aggression' (Butler 2009: 167). Each of these possible routes to action starts with the same state of being 'injured and rageful'. What distinguishes them is the act that follows, whether it is using one's own hurt to justify one's hurtfulness, or whether it is struggling towards 'non-violence' as a consequence of knowing, deeply, from one's own experience, how damaging violence can be. Given that Butler has spoken about what she calls 'pre-emptive reparation' and that she has engaged explicitly with Klein in her work on violence (Butler 2009), this analysis can be taken to have a Kleinian set of resonances. The radical move here, politically as well as psychosocially, is to use the insight that there could be a fundamental impulse towards care arising from what Butler names as 'a more general conception of the human … in which we are, from the start, given over to the other' (Butler 2004: 31) to generate an impulse towards reparative and generous reaching out towards others, what elsewhere might be called relationships of trust.

As noted earlier, while the articulation of a culture of care can start with destructiveness and move towards reparation, it can also be expressed differently, as a consequence of a different strand of thinking, also psychoanalytic in formation, about what is needed for a benign sociality. I have discussed this previously under the heading of 'relational ethics' (Frosh 2011) that stem in large part from the object relational perspective in psychoanalysis, or more generally, the 'relational' one that de-emphasizes drives and instead focuses on the intersubjective conditions under which people develop and then go on to live their lives. The varying versions of this, particularly associated with Donald Winnicott but latterly with a large group of significant American psychoanalysts (including Stephen Mitchell, Jessica Benjamin, and Thomas Ogden), concern how the 'inner worlds' of human subjects are formed as a process of internalization and identification with loved 'objects'; and the conditions under which this happens are given as products of social structures. That is, the social conditions under which people live profoundly influence their object relationships and their inner structures of security and

selfhood. Briefly, we can take as an example the sociologist Axel Honneth (1996), for whom 'affective recognition' grounded in childhood experiences in the family is the source of the kind of emotional stability and security of selfhood that is necessary for social life. It is also a term employed by psychoanalytic theorists who wish for equality and benevolence to operate between analysts and patients. Much of what is discussed under this heading concerns the questions of what constitutes recognition and whether it can ever be enough to sustain an ethical relationship. In social theory, recognition is one focus of a debate about equality that addresses identity politics, including those surrounding gender, race, and sexuality. Without recognition, identity cannot be manifested in an emancipatory way; groups are disowned, sidelined, stigmatized. Recognition politics demands that a space is made for these stigmatized groups and that they are acknowledged as a site of existence and value. But it also emphasizes the importance of being recognized *by the other*: it is this that creates the conditions for the development of a secure self that can act in the world, that can be an *agent* as well as a 'subject', much in line with the theory of self-formation put forward by Winnicott (1975), on which several of these thinkers explicitly draw. As Benjamin (2000: 294) comments, 'The problem of whether or not we are able to recognize the other person as outside, not the sum of, our projections or the mere object of need and still feel recognized by her or him, is defining for intersubjectivity'.

There are many criticisms that can be made of this general relational approach (Frosh 2010), but what I want to emphasize here is not so much the strengths or weaknesses of the various theories, but rather that from the nuanced and sophisticated account of subject formation developed by different strands of psychoanalysis (here, Kleinian and object relational), different but not incompatible ideas about social democracy can arise. These focus on the conditions under which what Butler (2004) calls 'precarious lives' can be made more secure. That is, from their somewhat distinct vantage points, these psychoanalytic ideas move from a sensitive awareness of the vulnerability of human subjects to considerations of how specific modes of social organization can moderate or produce states of increased vulnerability or greater security, exaggerated violence or more resilient possibilities of non-violent, reparative practice. None of this is revolutionary in the standard sense; but it is part of the struggle for social conditions that will be better attuned to the preservation and furtherance of less precarious lives.

Political Dimensions of Desire

The terms of this discussion to date have left some very obvious gaps. One is the absence of Lacanian theory in what has been outlined earlier, yet it is in relation to the 'Lacanian left' (Stavrakakis 2007) that some of the strongest contemporary appeals to radical thinking can be found.[1] There are many ways into this, but one issue of concern is that while Lacanianism seems to have a strong political analysis born out of its interest in the functioning of the Symbolic order of language and culture, its actual engagement with

politics is at a level of such abstraction that its effect can be conservative rather than truly progressive. To unpack this a little, the division Lacanianism makes between different registers of experience—Imaginary, Symbolic, Real—has proved very fertile for thinking about the way in which the human subject is 'subjected to' the operations of a social order that regulates it and makes demands of it, yet also retains enough difference (agency, perhaps, in terms that do not quite belong to the Lacanian domain) to resist that order. The problem is that this 'resistance' is not in itself a politically active stance. Instead, it is built out of an acknowledgement of 'lack', understood in the Lacanian scheme as linked inextricably to desire. This is a complex notion that need not delay us too much here. The key point is that desire is fuelled by the lack *of* the object that will satisfy that desire (the breast, for instance, which can meet the infant's 'need' but never fully sate the 'demand' for love) and also by the lack *in* the 'Big Other', which might be approximated to by the Symbolic order or by the fantasy relation between the subject and 'society' as a whole (i.e. the Big Other is not the same as 'society', but refers to the subject-social relationship, mediated as that is by fantasy). The outcome of this is that the subject experiences a lack in itself, *and* in the social world. Lacan's clearest description of this situation in a way that has political relevance comes from his Seminar XI:

> A lack is encountered by the subject in the Other, in the very intimation that the Other makes to him by his discourse. In the intervals of the discourse of the Other, there emerges in the experience of the child something … namely, *He is saying this to me, but what does he want?* The desire of the Other is apprehended by the subject … in the lacks of the discourse of the Other, and all the child's *whys* reveal not so much an avidity for the reason of things, as a testing of the adult, a *Why are you telling me this?* ever-resuscitated from its base, which is the enigma of the adult's desire

> (Lacan 1973: 214).

This seems to be a commentary on how the subject comes to be called into being by a social order that 'desires' it. The subject is lacking because it is placed in a symbolic web that cuts it off from the source of its desire; that is, the subject is always constructed from 'outside' in the light of the desire of the Other—the mass of unconscious expectations and wishes that are directed towards it and position it as one subject amongst others. But the Other is also lacking, because it makes demands on the subject, because it clearly desires that the subject plays a particular role or occupies a specified position (for instance, as a consumer or loyal citizen). This faces the subject with the question, 'What does the Other want?'

The interesting point here is that the Other is conceptualized as having a lack within itself that produces a desire of its own, which the subject is immersed in. This suggests that any society is propped up by the flow of fantasies that come from its subjects; that is, a social order requires a kind of *investment* from its subjects that can maintain it. Derek Hook (2008) offers as an example of this how the perpetuation of apartheid in South Africa depended both on a set of fantasy investments by its proponents, fuelled by

envious longings and anxieties as well as greed, *and* on the existence of a social order that produced absence and exclusion as a means of sustaining itself, and hence inevitably generated a field of desire. Apartheid, which was clearly a structural (economic and political) organization of society along highly racialized, colonial lines, survived not only on violence, but also because of the intense emotional dependence on it by a White population that dealt with its own lack through anxious investments in privilege and projective fantasies of black 'primitivity' and hatred, and through (as Hook documents) revulsion towards mixing. Hook (2008) provides an example here, derived from J. M. Coetzee's (1991) analysis of the writings of an 'apartheid ideologue, Geoffrey Cronjé':

> any given fantasy is divided between its beatific, stabilizing aspect (the promise, say, of absolute white racial purity/superiority) and its vexing, radically destabilizing aspect that forms the basis of a variety of exaggerated threats (for Cronjé, the contaminant of blackness, or, metonymically, the danger of infection by black blood). The dynamic interplay of these aspects should not be lost on us: the beatific dimension of fantasy functions to mask a structural impossibility (a pure, independent Afrikaner community existing in a state of self-contained harmony), whereas the second dimension provides the reason—and typically also a scapegoat —for why such an inherent impossibility could not be realized

> (Hook 2008: 293).

More generally, the racist imaginary is clearly located at a social level—racism is a social phenomenon, in other words—yet it is deeply invested in by individuals and is often sustained by its immense affective load. Indeed, whereas Kleinian theory has offered at times vivid accounts of how racism might operate through processes of projection of split-off 'bad' material into socially nominated denigrated others, Lacanian psychoanalysis' contribution to understanding and contesting racism lies mainly in unpicking the way racism functions to cover over certain kinds of lack at the level of the individual, for instance through fantasies of the racialized other's 'theft' of the possibility of the subject's full 'enjoyment' (sexuality, wholeness, integrity, success, etc.), as well as at the level of society (migrants as scapegoats for economic failure, etc.) (Frosh 2013; Hook 2012).

There are some contradictory political implications of the social order's need for recognition by the subject. On the one hand, this account shows how the subject is caught in a web of desire that means it is always left empty of something, because the Big Other is not capable of delivering on its promises. In fact, capitalism relies on this non-delivery: continuing consumption depends on raising the hope of fulfilment and then always failing to meet this hope, while holding open the possibility that further consumption will finally make one fully satisfied. On the other hand, the incompleteness of the Big Other, its constitutive lack, allows for the possibility of some resistance as the subject discovers its own unmet desire. Stavrakakis (2007) frames it like this:

> In fact it is this constitutive and unbridgeable gap between the symbolic/imaginary nexus (the field of social construction and institution) and the always escaping real which also makes history possible: if it was feasible for a particular social construction to symbolise fully the real, then history would come to an end, together with the permanent play between human creativity (desire) and social dislocation (lack) (2007: 40).

Stavrakakis presses this point to argue that the Real—the dimension of experience that can never be properly symbolized—acts as a kind of pressure for change. That which is left out of symbolization always seeks to return, creating new possibilities and perhaps opportunities for the oppressed to have their 'voices' heard.

Stavrakakis (2007), however, also notes another feature of the typical Lacanian political scene, characterized in his account by the work of Slavoj Žižek. This relates to the argument that the subject can never cease to be split, because once the subject achieves desire (i.e. ceases to be lacking) it comes face to face with the horror of its true dimension of need. It is unclear at times whether this is simply a logical statement (desire is defined in relation to what is lacking—the *object a*, as Lacanians name it—hence, if the object is achieved and is no longer lacking, it cannot be desired) or whether it is a description of a psychoanalytic observation (we never seem to desire what we have). In either case, it is not clear that the 'no lack = no desire' formula *necessarily* holds, even if it is often true: is it *always* the case that marriage kills desire? Are there no desires left once one has been fulfilled? This suggests that the assumption made by some Lacanians that the achievement of desire would be terrifying because it would result in the collapse of the subject as a subject of desire, may also be driven more by the semantic logic of the 'no lack = no desire' formula (perfect achievement of desire is impossible) than by a study of what actually happens. Jason Glynos (2001), for instance, claims that:

> what is most traumatic is not that I am subject to the rule of the big Other, to the Master. All our complaints and appeals to justice conceal their true function, namely to *maintain* the big Other and the *jouissance* it makes possible for us. Far more traumatic is the possibility that the big Other does not exist. *This* is ultimately what we cannot accept as subjects of desire and this is ultimately the reason for our ready recourse to fantasies of the 'Other of the Other' who 'steal' our enjoyment (2001: 97).

In some ways this is a persuasive idea, articulating how an 'administered' society might promote anxiety in its subjects ('where is the Big Other who will ensure my needs are met?') which then is dealt with through projection into others who can be elected as the derogated sources of psychic suffering, a role that has been filled particularly by 'migrants' in recent times. It also constitutes a kind of political theory claiming that authority functions to channel resistance ('the existence of the Big Other frustrates my desires, maintaining them and giving sense to my life') and so sustains the security of the subject who never has to face the problem of what to do with a desire that has been fulfilled. What it also makes problematic, however, is the possibility of political resistance

or revolution: if a masterful Big Other is always needed in order to preserve the subjecthood of the desiring subject, how can change ever occur? We can be left here with nihilism or ungrounded messianism, in which the only way to produce political change is to smash the social system completely, under the guise of a breakthrough of the Real; and as this is intolerable, because it would also mean the demise of the subject, it means that we are caught politically between a rock and a hard place.

> For Lacanians, the return of the Real is always a disruptive, almost revolutionary event which shatters the entire social totality constructed around its exclusion. Every social order, therefore, has a single touchy 'nodal point' which it must maintain, or else it will collapse. Since the exclusion of a Real element is supposed to be necessary, Lacanians urge that one reconcile oneself to the inevitability of lack. Lacanian politics is therefore about coming to terms with violence, exclusion and antagonism, not about resolving or removing these
>
> (Robinson 2004: 260).

While this might be an overly critical statement, it reflects a difficult problem that Žižek has also referenced directly, for example when discussing Lacan's idea that the position of the psychoanalyst (the 'discourse of the Analyst') is one that, ideally, disrupts the lure of absolute authority (the 'discourse of the Master') and returns power to the subject. 'Lacan's claim', Žižek writes (2006), is:

> that the discourse of the Analyst prepares the way for a new Master ... The question, however, remains: how, *structurally*, does this new Master differ from the previous, overthrown one ...? If there is no structural difference, then we are back with the resigned conservative wisdom about (a political) revolution as a revolution in the astronomic sense of the circular movement which brings us back to the starting point (2006: 307).

One response to this is to see the overthrowing of the Master as only the first step towards building a new set of structures, but the nature of these structures remains hazy, in contrast to the relatively modest proposals of some of the other theorists described in this chapter, who seem to think that it might be possible to use psychoanalytic ideas to promote a more progressive, more caring, and relationally fulfilling society. Still, there is something alluring about the Lacanian emphasis on lack and on constant critique: amongst other things, it describes a restlessness that is at the heart of psychoanalysis when it comes to confronting the rigid structures of power and the assumed norms of what Freud, scathingly, called 'civilization'.

References

Adorno, T. (1967). 'Sociology and Psychology, Part I and II'. *New Left Review 46*: 67–97.

Bailly, L (2009). *Lacan: A Beginner's Guide*. Oxford: Oneworld.

Benjamin, J. (2000). 'Response to Commentaries by Mitchell and by Butler'. *Studies in Gender and Sexuality 1*: 291–308.

Brickman, C. (2003). *Aboriginal Populations in the Mind*. New York: Columbia University Press.

Butler, J. (2004). *Precarious Life*. London: Verso.

Butler, J. (2009). *Frames of War*. London: Verso.

Butler, J. (2015). *Notes Toward a Performative Theory of Assembly*. Cambridge, MA: Harvard University Press.

Coetzee, J. (1991). 'The Mind of Apartheid: Geoffrey Cronjé'. *Social Dynamics 17*: 1–35.

Danto, E. (2005). *Freud's Free Clinics: Psychoanalysis & Social Justice*. New York: Columbia University Press.

Einstein, A. and Freud, S. (1933). 'Why War?'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 22 (pp. 195–216). London: Hogarth Press.

Freud, S. (1908) '"Civilized" Sexual Morality and Modern Nervous Illness'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 9 (pp. 177–204). London: Hogarth Press.

Freud, S. (1913) 'Totem and Taboo'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 13 (pp. vii–162). London: Hogarth Press.

Freud, S. (1919) 'Lines of Advance in Psycho-Analytic Therapy'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 17 (pp. 157–168). London: Hogarth Press.

Freud, S. (1920) 'Beyond the Pleasure Principle'. In J. Strachey (ed and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 18 (pp. 1–64). London: Hogarth Press.

Freud, S. (1927). 'The Future of an Illusion'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 21 (pp.1–56). London: Hogarth Press.

Freud, S. (1930). 'Civilization and its discontents'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 21 (pp. 57–146). London: Hogarth Press.

Freud, S. (1939) 'Moses and Monotheism'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 23 (pp. 1–138). London: Hogarth Press.

Freud, S. (1961). *Letters of Sigmund Freud 1873–1939*, ed. E. Freud. London: Hogarth Press.

Frosh, S. (1989). *Psychoanalysis and Psychology: Minding the Gap*. London: Macmillan.

Frosh, S. (2005). *Hate and the Jewish Science: Anti-Semitism, Nazism and Psychoanalysis*. London: Palgrave.

Frosh, S. (2010). *Psychoanalysis outside the Clinic: Interventions in Psychosocial Studies*. London: Palgrave.

Frosh, S. (2011). 'The Relational Ethics of Conflict and Identity'. *Psychoanalysis, Culture and Society 16*: 225–243.

Frosh, S. (2013). 'Psychoanalysis, Colonialism, Racism'. *Journal of Theoretical and Philosophical Psychology 33*: 141–154.

Frosh, S. (2017). 'Primitivity and Violence: Traces of the Unconscious in Psychoanalysis'. *Journal of Theoretical and Philosophical Psychology 37*: 34–47.

Giffney, N. and Watson, E. (eds) (2017). *Clinical Encounters in Sexuality: Psychoanalytic Practice and Queer Theory*. New York: Punctum.

Glynos, J. (2001). '"There is no Other of the Other" Symptoms of a Decline in Symbolic Faith, or, Žižek's Anti-Capitalism'. *Paragraph 24*: 78–110.

Goggin, J. and Goggin, E. (2001). *Death of a 'Jewish Science': Psychoanalysis in the Third Reich*. West Lafayette, IN: Purdue University Press.

Goldner, V. (2003). 'Ironic Gender/Authentic Sex'. *Studies in Gender and Sexuality 4*: 113–139.

Hinshelwood, R. (1994). *Clinical Klein*. London: Free Association Books.

Hollway, W. (2006). *The Capacity to Care*. London: Routledge.

Honneth, A. (1996). *The Struggle for Recognition: The Moral Grammar of Social Conflicts*. Oxford: Polity.

Hook, D. (2008). 'Fantasmatic Transactions: On the persistence of apartheid ideology'. *Subjectivity 24*: 275–297.

Hook, D. (2012). *A Critical Psychology of the Post-Colonial: The Mind of Apartheid*. London: Routledge.

Isin, E. (2004). 'The Neurotic Citizen'. *Citizenship Studies 8*: 217–235.

Jacoby, R. (1975). *Social Amnesia*. Sussex: Harvester.

Jacoby, R. (1983). *The Repression of Psychoanalysis*. New York: Basic Books.

Khanna, R. (2004). *Dark Continents: Psychoanalysis and Colonialism*. Durham, NC: Duke University Press.

Klein, M. (1935). 'A Contribution to the Psychogenesis of Manic-Depressive States'. *International Journal of Psycho-Analysis 16*: 145–174.

Klein, M. (1957/1975). 'Envy and Gratitude'. In M. Klein, *Envy and Gratitude and Other Works*. New York: Delta.

Lacan, J. (1973/1979). *The Four Fundamental Concepts of Psychoanalysis*. Harmondsworth: Penguin.

Laplanche, J. and Pontalis, J.-B. (1967/1973). *The Language of Psychoanalysis*. London: Hogarth Press.

Makari, G. (2008). *Revolution in Mind*. London: Duckworth.

Marcuse, H. (1955). *Eros and Civilization*. Boston, MA: Beacon Press.

Reich, W. (1946). *The Mass Psychology of Fascism*. Harmondsworth: Penguin.

Rieff, P. (1966). *The Triumph of the Therapeutic*. Harmondsworth: Penguin.

Robinson, A. (2004). 'The Politics of Lack'. *British Journal of Politics and International Relations 6*: 259–269.

Roudinesco, E. (2016). *Freud: In His Time and Ours*. Cambridge, MA: Harvard University Press.

Rubin, A., Mandelbaum, B., and Frosh, S. (2016). '"No memory, no desire": Psychoanalysis in Brazil during Repressive Times'. *Psychoanalysis and History 18*: 93–118.

Rustin, M. J. (1991). *The Good Society and the Inner World*. London: Verso.

Rustin, M. J. (1995). 'Lacan, Klein and Politics: The Positive and Negative in Psychoanalytic Thought'. In A. Elliott and S. Frosh (eds), *Psychoanalysis in Contexts*. London: Routledge.

Ryan, J. (2017). *Class and Psychoanalysis*. London: Routledge.

Segal, H. (1990). *Dream, Phantasy and Art*. London: Routledge.

Stavrakakis, Y. (2007). *The Lacanian Left*. Albany, NY: State University of New York Press.

Stonebridge, L. (1998). *The Destructive Element: British Psychoanalysis and Modernism*. London: Macmillan.

Traverso, E. (2016). *The End of Jewish Modernity*. London: Pluto Press.

Winnicott, D. W. (1965). *The Maturational Processes and the Facilitating Environment: Studies in the Theory of Emotional Development*. London: Hogarth Press.

Winnicott, D. W. (1975). *Through Paediatrics to Psycho-Analysis*. London: Hogarth Press.

Zaretsky, E. (2015). *Political Freud*. New York: Columbia University Press.

Žižek, S. (2006). *The Parallax View*. Cambridge, MA: MIT.

Notes:

(1.) For an accessible general introduction to Lacanian theory, see Bailly (2009).

Stephen Frosh

Department of Psychosocial Studies, Birkbeck, University of London

Oxford Handbooks Online

Abstract and Keywords

This chapter develops a concept of epistemic anxiety and explores its particular relevance to educational settings and to wider contexts in contemporary society. The idea brings into conjunction two significant psychoanalytical ideas. The first is that of the epistemophilic instinct proposed, following Freud's reflections on anxiety, by Melanie Klein and developed by Wilfred Bion into a theory of thinking and its relations to love and hate. The second is the theory of unconscious social defences against anxiety which first evolved within the Tavistock tradition of psychoanalytic social research in the 1950s and has recently been subject to reappraisal and new applications. Its argument is that epistemic anxiety frequently arises in learning situations, and that learning and thinking can be facilitated if this is understood.

Keywords: epistemophilic instinct, unconscious anxiety, social defences, learning difficulties, Melanie Klein, Wilfred Bion

Introduction

This chapter proposes that a specific kind of anxiety, identified as epistemic anxiety, can arise from the desire or impulse to know. It suggests that this form of anxiety is particularly significant in those spheres of social life which are concerned with learning and knowledge, for example in educational settings, and that it may also have a wider presence in contemporary society. The concept may also have relevance to the clinical context of psychoanalysis, but this is not the main subject of this chapter.

This formulation has come about through bringing into relation with one another two fields of psychoanalytic theory—the first concerned with unconscious anxiety, and the second with the 'epistemophilic instinct'—in the specific context of the 'British school' and especially its filiation through the writings of Freud, Klein, and Bion. It is from this

conjunction that it became possible to identify anxiety about learning and knowing as a specific kind of anxiety.

The concept of unconscious anxiety and its origins and objects originates in Sigmund Freud's work and was further developed in the writings of Melanie Klein. The investigation and elaboration of the forms and functions of anxiety have been of great importance in the development of psychoanalytic theory and clinical practice since the beginnings of psychoanalysis, and have continued to evolve. The idea of defences against anxiety has also found significant applications to social institutions and wider social contexts.

The second field concerns the idea of an epistemophilic instinct, or an innate desire for knowledge. This also has its source in Freud's work, was then developed by Klein, and became more fully established as a central postulate of psychoanalytic theory in the writings of Wilfred Bion, through his argument that the impulse to know is as important to the mind and its development as the impulses to love and to hate. This gave rise to Bion's formulation of L, H, and K (a kind of algebraic triad) as the central constituents or dispositions of the mind. In recent years, this focus on the vicissitudes of the disposition to know (in Bion's terms, 'K', 'no K', and 'minus K') has become a major theme in the development of psychoanalytic theory and practice.

Anxieties and Their Objects

Freud wrote about anxiety most extensively in *Inhibitions, Symptoms and Anxiety* (1926). The principal anxiety with which he was concerned was the anxiety associated with the Oedipus complex, which he formulated as castration anxiety. This was aroused by fears of paternal retribution fantasized to be the consequence of the infant's sexual desire for the mother. Freud believed that this found expression in obsessional neurosis, including obsessional phobias, whose objects functioned unconsciously as condensations or displacements of Oedipal anxieties. Freud interpreted Little Hans's phobia of horses in terms of his anxieties about his father (Freud 1909a), and the Rat Man's phobias about torture by rats also within the context of Oedipal anxieties about being punished for his sexual feelings by his father, and by concern about what might happen to his father and his fiancée (Freud 1909b). Freud also referred to anxieties about object loss, but this had a subordinate place in his reflections on anxiety in the 1926 paper. The theme of responses to object loss was given its fullest development in *Mourning and Melancholia* (1917).

Klein's theories of anxiety reflect the significant differences that emerged between her ideas and Freud's. She proposed that there were two primary kinds of anxiety, those associated with the depressive and paranoid schizoid positions, respectively (Klein 1940, 1946). The central issue in the generation of anxiety arose from the self's relation to its objects, and the place of love and hate in that relation. In the paranoid–schizoid position, love and hatred were deeply split from one another, both within the self and in its relation to its objects. Love was directed towards the self and those objects with whom it

identified, hatred towards others who were defined in negative terms. Feelings of hatred directed towards split-off objects gave rise to anxieties about the damage which those objects might inflict in reprisal for the aggressive feelings directed towards them. Paranoid–schizoid states are associated with a harsh superego, and a belief that 'badness' must be punished wherever it lies, inside or outside the self. Here the talion principle—an eye for an eye and a tooth for a tooth—rules in the sphere of justice.

The second kind of anxiety arises from the capacity for concern for the state of the object, as the recipient of the states of mind and feelings of the self towards it. Anxiety is thus aroused not only by threats to, and fears for, the state of the self, but also by fears for the state of the other. Depressive anxiety is thus closely linked to guilt, and gives rise to feelings of remorse.

Klein believed that both the dispositions to hate and to love were innate. She thought feelings of hate arose in infants in response to the frustrations inevitable in early life, and from emotions of envy and rage which could be provoked by dependence on a mother who did not and could not always provide what the baby desired. She believed that Oedipal issues—anxieties about mother's relationship with father and about rival siblings—arose much earlier than in Freud's account of early life. It depended both on the actual environment of parental love and care, and on the infant's innate vulnerability and disposition, what balance between love and hating impulses would be the outcome of internal conflicts in the infant's mind. Klein gave particular emphasis to the infant's inner world, rather than to deficits in its external environment, which she felt were more widely recognized at that time. Both paranoid–schizoid and depressive anxieties (about the damage inflicted by the infant in phantasy on its loved objects) were potent sources of anxiety in her view. This emphasis in her work on the internal world, and on destructive feelings, gave rise to differences between the Kleinian school and perspectives such as that of attachment theory, and the Independent tradition of psychoanalysis which attributed a larger role to environmental deficits in causing developmental difficulties. Bion's theory of the functions of 'containment' and the relations of 'container and contained', although located within the Kleinian tradition, has rebalanced this debate by focusing on the unconscious dimensions of the mother's contribution to the infant's emotional development.

Klein believed that hatred is often in any case the outcome of disappointment or disillusionment with loved objects. The infantile origin of a criminal state of mind could in her view lie in the hatred and feelings of persecution aroused by loved objects which were felt unconsciously to have betrayed the self or to have been damaged by it. Because she thought that such hatred was originally directed towards loved objects, it followed that its interpretation and understanding, for example through psychoanalysis, could reveal an underlying capacity for love. In fact, idealization makes some element of disappointment inevitable for everyone. This led Klein to be hopeful about the psychoanalytic treatment of children with 'criminal tendencies' (Klein 1927). Because in depressive states, the good and bad, the loved and the hated, are not so deeply split off from one another, these states may make possible, she thought, some interest in and

concern for the perpetrators of bad deeds, as well as for their victims. Forms of justice which involve recognition and reparation, as well as vilification and punishment, become conceivable. The idea that the self might be assailed by feelings of concern for the harms it has caused to others, rather than only by anxiety about threats to itself, was a crucial contribution by Klein to the understanding of anxiety.

Although Klein believed that the onset of the Oedipus complex occurred in the first year of life, earlier than Freud thought, a good deal of the discussion of the paranoid–schizoid and depressive positions has referred to dyadic rather than triadic relationships. The focus has tended to be on relations between the self and its primary object, whether these take the form of paranoid–schizoid splitting or evolve into depressive concern. This emphasis reflected a 'maternal' emphasis in one phase of development of object-relations psychoanalysis, in its Winnicottian as well as Kleinian variants. But there has been a renewal of interest in Oedipal issues in recent years (a return to psychoanalysis of the repressed father figure, one might say) and this has made it easier to relate both paranoid–schizoid and depressive anxieties to the early Oedipal situation. In Klein's view, this is a central locus of both paranoid–schizoid feelings (liable to violently split the parental couple in phantasy), and depressive ones, where it is possible for love to dominate in the relations with both parental objects, outweighing feelings of jealousy or envy towards them. In such a state, the separation of parental objects from the infantile self, and their closeness to each other, can be tolerated and even enjoyed, if there is, as it were, enough love to go round. This renewal of interest in the Oedipal situation has had considerable importance in the growth of understanding of the epistemic issues which will be considered later in the chapter.

Psychoanalytic investigations of autistic and psychotic states in children, by analysts who include Frances Tustin (1972), Donald Meltzer et al. (1975), and Thomas Ogden (1992), have given rise to the postulation of a stage of personality development which precedes the emergence of splitting and the paranoid–schizoid position. Ogden named this earliest state of mind the autistic–contiguous position, which he suggests is one in which a distinction between self and object has yet to be made. The refusal of object relations in autistic states, or their explosive or implosive disintegration in psychotic states, are the pathologies which correspond to this state. A particular kind of anxiety—annihilation anxiety—may be held to correspond to this state of mind or 'non-mind'. Such anxiety expresses the terror of complete disintegration and non-existence. It is not clear whether this is best understood as a normal first stage of infant development, or whether it rather corresponds to a state of mental collapse which can occur when relations with objects are felt to have entirely failed. The state of bodily and mental 'falling to pieces', which Esther Bick (1968) has described in certain infants, and the experience of 'catastrophic change' described by Bion (1970) (which may have the creative outcome of reintegration as well as giving rise to the psychotic collapse of the mind) seem close to the state which Ogden describes. W. Fred Alford (1997) has suggested that the kinds of evil of which mass exterminations are an example may be understood as projective responses to anxieties about annihilation.

Epistemic Anxiety

Psychoanalytic theories of anxiety have found a fairly extensive application to the spheres of group and institutional life. It is the experience of anxiety which is held in Bion's (1961) work on group processes to stimulate the oscillation found in groups between 'basic assumption' behaviour and 'work group' behaviour. In the former, the group's relations to its objects take the form of unconscious phantasies which relieve anxiety states. In the latter, the reality principle is in the ascendancy, and the real objects of work-related tasks and functions are present in the minds of the group. Bion identified three varieties of 'basic assumptions'. These are 'fight–flight', in which the group unites against a phantasied enemy; 'pairing', in which the group evolves the phantasy that a pair will together produce something extraordinary which can resolve its anxieties and sufferings: and 'dependency', in which the group is passive while expecting its leader to resolve these. Bion developed these ideas through his pioneering form of group therapy, and they have become widely influential in that field and in the larger educational practice of group relations. David Armstrong (2007: 139–50) holds that Bion maintained that the rational functioning of the 'work group' most of the time prevailed over irrational phantasy, although this positive view is not usually ascribed to Bion.

In Elliott Jaques's (1951, 1955) and Isobel Menzies's (1960) (later Menzies Lyth 1988) writings about social defences against anxiety, the topics for investigation were anxieties generated within contexts of work. In Jaques's case, anxieties were focused in the relations between managers, workers, and their trade union representatives, and were of both paranoid–schizoid and depressive kinds. In Menzies's work, the primary source of anxiety was the mental pain evoked by the actual tasks of nursing, arising from the proximity to suffering bodies, and the emotional distress which these evoked in trainees. She was less specific about whether these anxieties were paranoid–schizoid or depressive in character. However, it seems they were of both kinds, with depressive anxieties aroused by contact with the patients, and paranoid–schizoid or persecutory ones aroused by the authoritarian approach of the hospital. (William Halton (2015) has suggested that the obsessional nature of the nursing system reflected an element of unconscious hatred of the patients).

Each of these socio-psychoanalytic researchers described characteristic modes of unconscious defence which had been constructed by institutions and their cultures to manage anxiety states. This field of investigation has recently been extended (Armstrong and Rustin 2015) in reflections on changed theoretical understandings of anxiety, changing conditions in society, and by exploration of anxieties and defences which arise in different occupational settings. The Kleinian psychoanalytic concepts of anxiety which informs this work had as its primary locus, as we have seen, states of mind characteristic of early infant development—the infantile parts of the self—which become visible and accessible in the transference relationships of the clinical consulting room, and which are manifested, as Freud also believed, in the symptoms of mental illnesses. In adult life, these anxieties can arise within institutional and social contexts, in relations between the self and the persons and situations which it encounters. The question of whether psychoanalytical ideas like these can be useful in the explanation of what happens in ordinary social settings is one which will be relevant to the main topic of this chapter,

epistemic anxiety, which it will be suggested can be aroused by the demands placed on the self to learn.

What virtually all of these discussions of anxiety share is the view that anxiety is essentially a response to perceived danger to the self, or to its loved objects, whether this be at a conscious or unconscious level. Anxiety can in favourable circumstances mobilize the 'reality principle', as Freud named it, i.e. the capacity to respond rationally to perceived risks to the self or its objects. But in unfavourable circumstances, responses to anxiety can be unconscious and irrational, and can take the mind further away from an engagement with reality, rather than closer to it. Both Freud (1926) and Klein (1930) were at pains to point out that anxiety was a necessary stimulus to growth and development, as well as a source of danger to the self and its psychic equilibrium. Klein's argument was that anxiety is provoked by the infant's sadism towards its maternal object, and that this anxiety stimulated the developments of phantasy and symbol formation which could then enable it to be overcome. She described this process in relation to her analysis of a child patient, Dick. There can be differing views, perhaps specific to national cultures at a given time, of how much and what kinds of anxiety are excessive and damaging rather than healthy and motivating.

One can relate these views to the 'basic assumptions' (or structures of phantasy) first described by Bion in his Experiences in Groups (1961). Some have argued that 'dependency cultures', such as that of the British welfare state and some of its component institutions, enact an aversion to the anxieties which follow from competition and individual self-reliance, and constitute defences against these (Khaleelee and Miller 1985; Khaleelee 2003). 'Fight–flight' cultures, such as that found in a highly competitive market society such as the United States, are by contrast hostile to dependency, and intolerant to those perceived to have failed. In such societies, systems of social protection are much weaker. Social systems can ascribe different value to security and freedom.

The Epistemophilic Instinct

The origins of the idea of an epistemophilic instinct lie in Freud's writings about infantile sexual curiosity which are a central theme of his case studies of 'Little Hans' (1909a) and the 'Rat Man' (1909b). Freud describes Little Hans's keen interest both in the anatomical differences between little boys and little girls, and in the question of how babies are made. The Oedipal situation and its attendant anxieties (not least castration anxiety) are Freud's central topic in these two case studies, and indeed in his theory of infantile development. The term 'epistemophilic instinct' as such appears for the first time in the case study of the 'Rat Man'. Freud there argued that:

> the epistemophilic instinct is a preponderant feature in the constitution of an obsessional patient … The thought process becomes sexualized, for the sexual pleasure which is normally attached to the content of thought becomes shifted on

to the act of thinking itself, and the satisfaction derived from reaching the conclusion of a line of thought is experienced as a *sexual* satisfaction (1909b).

The epistemophilic instinct first contributes to the substitution of action by thought, and then 'procrastination in *action* is replaced by lingering over *thoughts*' (1909b: 245). Freud uses the term again in his 'Introductory Lectures' (Freud 1917b: 327–8).

The concept of an epistemophilic instinct becomes more fully elaborated in three of Melanie Klein's papers (Klein 1928, 1930, 1931). Her development of this idea follows Freud's focus on the Oedipal preoccupations of children but, because she locates these at an earlier stage of development than Freud, these states of mind are described as taking a more primitive form. The epistemophilic instinct in its earliest manifestations in Klein's work is conjoined with powerful oral and anal impulses directed aggressively towards mother's body, as well as with Oedipal anxieties about the presence of father's penis and other babies inside it. The epistemophilic instinct, or the desire for knowledge, may be inhibited or disrupted, Klein believed, if these phantasies are too destructive in their content. In one of her cases, she prefigured the later psychoanalytic idea that such phantasies of damage and destruction may be a cause of the severe inhibition of thought found in autistic states. Klein wrote of her patient Dick:

> his defences against his destructive impulses proved to be a fundamental impediment to his development. He was absolutely incapable of any act of aggression, and the basis of this incapacity was clearly indicated at a very early period in his refusal to bite up food. At four years old he could not hold scissors, knives or tools, and was remarkably clumsy in all his movements. The defence against the sadistic impulses directed against the mother's body and its contents … impulses connected with phantasies of coitus—had resulted in the cessation of his phantasies and the standstill of symbol–formation (1930: 224).

The fullest development of the idea of a primary desire to know or learn takes place, however, in the writings of Wilfred Bion (Bion 1962, 1970; Bléandonu 1994). He considered the construction of a mind or 'mental apparatus' in an infant to be a developmental achievement. This depends on nurture, on a maternal function of 'containment' through which the intense emotions of love and hate, and also the emotionally invested bodily sensations of the infant, are mentally processed by a mother or mother-figure, to a degree which maintains them at a tolerable level for the infant. Bion's more developed understanding of the maternal (in a functional sense) role in this process was influenced by Klein's idea of projective identification. This was her conception of the process whereby a baby projected unbearable aspects of its mental state into its mother for her to process on the infant's behalf. This had both an evacuative function (getting rid of what cannot be borne, as in an infant's screaming fit) but also one of communication. In normal circumstances a hungry baby will have an expectation, which is both innate and reinforced by experience, that crying is likely to bring a response from its mother.

Bion proposed that the impulses or dispositions to love, hate, and know were of equal significance in the functioning of the mind, in a theoretical development which has become very influential in psychoanalysis. His argument was that where the container–contained relationship could not cope with the intensities of projections of love and hate, damage could be caused to the mind. Much of his most important clinical work was with patients suffering from psychosis, and he interpreted psychotic states of mind by analogy with those of infants whose containing object was insufficient or absent. (Understanding of a psychotic state according to this conceptual model is distinct from a causal assertion that psychotic states in adults have this infantile origin, even though they may well often do so). The crucial developmental path for Bion was one where it becomes possible for primitive feelings to be transformed into thoughts, in a process which he named 'alpha function'. This might be thought of as a psychoanalytic formulation of the transformation of 'impressions' (sensory experiences) into ideas in the philosophy of the British empiricists. For Bion, however, the success of such transformations could by no means be taken for granted, and depended on favourable conditions of what he called 'containment'.

Hanna Segal (1957) argued, following Klein, that the capacity for symbol formation, or the function of 'K' in Bion's terms, depended on the attainment of the 'depressive position'. That is, on the lessening of paranoid–schizoid splitting to the extent that it was recognized that the apparently different objects of intense love and hatred were in reality the same objects, and that furthermore the love and hatred which were felt for these belonged to the same self. Thus the capacity for thought was one which depended on the understanding and integration of different emotions. Because for Bion the original objects of thought were the highly emotionally charged parental figures of the infant's mind, thinking itself was an experience from which emotions could not be excluded. 'Thinking' for Bion means something a little different from its ordinary meaning, in that he believes it to be an intrinsically emotional process. This is because it often involves the recognition or acceptance of something which cannot easily be recognized or acknowledged by the self beforehand. In other words, it involves a change in oneself, and not just a change in one's understanding of something, or in the contents of one's mind. This helps to explain later, as we shall see, why learning often involves anxiety. 'Learning from experience', which is an important concept in Bion's work, means learning from emotional experience.

This modern psychoanalytic interest in the nature and functions of 'knowing' has distanced itself from a dominant tendency in Western science and philosophy of sharply counterposing the domains of thought and feeling, cognition and the emotions. Indeed, it is commonly believed that they must be kept apart if knowledge is to be advanced, and the functions of the rational mind protected. Freud in his chosen role as a scientist for the most part worked within the framework of this distinction. Psychoanalysis in his view had discovered a new object for scientific investigation, the unconscious mind, but he believed that this could be understood by means of empirical evidence and logical arguments similar to those deployed by the major fields of science in which he intended psychoanalysis to find its place. Freud's purpose was to gain understanding of the unconscious by methods which would eliminate emotional biases or preconceptions, even

though one of his primary discoveries was to show how human beliefs and understandings of their experience were influenced by unconscious states of mind and feeling. The misperceptions and distortions of the symptom-dominated structures of the mind which he investigated—hysteria, obsessional neurosis, phobias, etc.—were states of mind driven by powerful emotions which functioned in an unrealistic or irrational relation to their objects. This did not call into question the nature of knowledge and understanding for Freud himself. His belief was that psychoanalysis could help both individuals and society more broadly to resist the unconscious and thus emotionally driven misrepresentations of experience, by understanding their origins at a deep level of the mind. But it nevertheless made emotions, conscious and unconscious, central to the understanding of how the mind actually worked, and challenges cognitivist approaches to psychology and the philosophy of mind.

Once psychoanalysts, from Klein onwards, began to investigate the earliest experiences of learning and understanding in infancy, emotions came even more to the fore in the understanding of mental function. This is because the initial object of curiosity—the desire for knowledge—is held to be the emotionally highly charged relationship of the infant with its mother. If the primary object of knowledge is, as Klein thought, the nature of mother's body and its inside, and the possible harm which the baby's hatred and bodily discharges may be doing to it, how can we explain the emergence of a cognitive function which is free from the influence of emotions?

The psychoanalytic argument in the work of Klein and Bion and their associates came to be that the functions of reason require a configuration of the mind within which aggression and anxiety are contained and limited. Where this does not happen, the capacities of the mind are liable to be fragmented or inhibited. Bion (1962) describes states of mental near-disintegration in his psychotic patients. The formation of the mind is held to depend on an experience of nurture in emotionally rich relationships in infancy and childhood. Segal (1957) described this as leading to the growth of the capacity for symbol formation. In this model of the development of the mind, the capacity for thinking itself depends on the integration of emotional states.

Bion's interest in the functions and processes of thought lead him to find some affinity between his work and Kant's philosophy of mind. In particular he draws a comparison between his own conception of an unknowable infinite to which the mind aspires, and Kant's distinction between the noumenal and phenomenal domains, only the latter being within the reach of our perceptual apparatus (Bion 1962: 67). But the distinctive assertion of psychoanalytic writers, particularly in the tradition of Klein and Bion, is that the development of the faculty of reason—Bion's domain of K—requires the recognition and integration of conflicting emotions, and is always at risk of disturbance from the resurgence of unconscious feelings. These include those intense forms of love and hatred which dominate paranoid–schizoid states of mind, and the anxiety, guilt, and remorse which are the emotions characteristic of Klein's depressive position.

Epistemic Anxiety

The two areas of psychoanalytic theory which have been outlined in the first parts of this chapter—that concerning anxiety and its varieties, and that concerning the different instincts characterized as those of love, hate, and curiosity, or the desire for knowledge (L, H, and K in Bion's formulation)—have seldom been brought into a direct relation to one another in the psychoanalytic literature. Theories of anxiety in Freud and Klein have assumed that the predominant drives which the personality needs to organize and reconcile in its encounters with internal and external reality are those of love and hatred, the libidinal and the aggressive impulses. The Oedipal anxieties described by Freud arise from the fears that the infant and child's libidinal desires will give rise to aggressive retribution which will threaten its survival, or at least the boy's or girl's sexual identities. The paranoid–schizoid and depressive anxieties conceptualized by Klein also arise from how the self organizes its dispositions to love and hate. In the paranoid–schizoid position, these are split off from one another, and the self feels assailed by persecutory anxiety. In the depressive position, the self becomes aware of its own aggressive impulses, and through its love for its object becomes aware of the harm that they can do to it. Neither Klein's nor Bion's addition of a third major instinct or impulse, curiosity, the epistemophilic instinct, or K, to the primary dispositions of love and hate, was accompanied by the identification or naming of a specific kind of anxiety, although one could see the states of mind of Bion's psychotic patients, with their massive projections and evacuations, as enactments of extreme anxiety about psychic survival and the fragmentation of the mind itself.

Bion writes about the experience of the 'rudimentary consciousness' of the infant when its projections of intense feelings cannot be received and processed by a maternal container, describing its feelings as 'nameless dread'. This is analogous to his description of the psychotic states of mind in his *Second Thoughts* (1962).

There seem grounds for proposing, however, that just as specific kinds of anxiety are linked closely to the instincts or dispositions of love and hate, so there may be a specific anxiety associated with the epistemophilic instinct, the desire to learn. It is this kind of anxiety that I term 'epistemic anxiety'. I suggest that this kind of anxiety is of common occurrence, parallel in this respect to paranoid–schizoid and depressive anxieties, and also needs to be differentiated from catastrophic forms of psychotic anxiety. Although in psychotic states the disintegration of the mind itself is a source of terror, it seems that in such states the impulses to love, hate, and to know are scarcely distinguishable from one another. The concept of 'epistemic anxiety' is here being used to denote a condition in which the mind and its disposition is certainly recognized to exist, but in which its capacities for thinking are felt to be inadequate to the demands of its internal and external environment.

The clinical consulting room and the transference–countertransference relationship between analyst and patient has been the primary locus of discovery of concepts and theories within psychoanalysis since Freud (Rustin 2009). The idea advanced by Bion and

by Edna O'Shaughnessy (1981) that the dimension of 'knowledge' or K can be specifically identified as an aspect of patients' communications, and their wish to discriminate this from preoccupations with love and hate, indeed suggests that we might identify a form of anxiety, and defences against anxiety, specific to this epistemic function within the transference situation of psychoanalysis itself. However, it proves challenging to separate epistemic anxiety about learning itself from other anxieties raised by the objects of thought and experience.

Descriptions of work with narcissistic or borderline patients by, for example, Herbert Rosenfeld (1987), Ronald Britton (1998, 2015), and John Steiner (1993) have certainly placed resistances to the understanding offered by analysts as a central clinical and theoretical issue. Rosenfeld wrote about resistance as the 'negative therapeutic reaction', and suggested that 'destructive narcissism' could have a large role in this. Destructive narcissism is the form of pathological personality organization which is dominated not by self-love, as in the usual concept of narcissism, but rather by unconscious identification with a superior, destructive, and envious part of the self. This part of the self resists interpretation and understanding, because of what this might reveal about the damage the self is inflicting on its loved objects. If the destructive forces which dominate the self can be recognized and understood, some relief may obtained, but the experience of being understood is felt to be a dangerous one, and flight from it often follows from phases of insight and improvement. Epistemic anxiety in this context arises from the extreme depressive and paranoid–schizoid anxieties to which an understanding of the self as dominated by hatred might lead.

Steiner took a central characteristic of borderline states to be the disavowal of knowledge, and the defence against knowing that he termed 'turning a blind eye'. He developed this idea in the context of an illuminating interpretation of Sophocles' play, *Oedipus the King*, revising Freud's interpretation of this play in the light of Bion's understanding of the mental function of K. Whereas Freud's emphasis was on the forbidden Oedipal desires which the play's tragic hero had unwittingly enacted, Steiner argued that decisive in the action of the play was the unwillingness of Oedipus and Jocasta to recognize what was hidden in plain sight, namely that Oedipus had actually fulfilled the prophecy that he would kill his father and marry his mother. In the situation represented in the play, the instincts of love and hate (the libidinal and aggressive impulses) and that of curiosity and its potential to be diverted by unconscious disavowal are deeply entangled with one another. The particular kind of knowledge which arouses intense resistance here is knowledge of particular objects, namely the Oedipal relationships between father, mother, and son, and the different ways in which these are enacted. (The infant Oedipus is, as his parents intended, sent to his death before he grows up to enact his part in what Freud saw as a primordial pattern of unconscious impulses).

Britton's (1998) analysis of different kinds of narcissistic personality organizations also gives a central place to the disposition to know or to refuse to know. His important contribution to the understanding of Bion's faculty of knowledge is the connection he

makes with the Oedipal situation. Britton's argument is that the emergence of the infant's capacity to learn and understand depends on its toleration of a 'third position' which makes possible the differentiation between reality and phantasy. Indeed, we commonly talk of 'triangulation' (deploying a metaphor whose origin lies in surveying) to describe reference to a third point of view as resolving differences which cannot be settled from two positions alone. Britton describes different ways in which the reality of the Oedipal situation, faced in phantasy by patients in relation to their analyst, can be denied. In an identification in phantasy with a maternal figure, emotional intimacy and oneness can be preserved, at the expense of the paternal function of thought. In an identification in phantasy with a paternal figure, a certain kind of thought becomes possible, but only if it is entirely cut off from emotion and intimacy.

Britton (1993) has given his argument a wider social dimension, in describing alternative modes of holding on to beliefs in defiance of reality. One of these involves an indissoluble attachment to words ('fundamentalism') and the other an equally intense attachment to images ('idolatry'), each involving an exclusive identification with a phantasy of a male or female parental function. The presumption of all these accounts is that some kind of meaningful ordering of experience, and belief in its true relation to reality, is essential to the self's sense of security and well-being. 'Epistemic disruption' is a painful and potentially catastrophic experience, both for individuals and groups.

In these accounts, and indeed in Freud's (1905, 1926) original discussions of the crucial role of patients' resistance in psychoanalysis, the refusal or disavowal of knowledge and understanding, and the inhibition of curiosity, has a large place, sometimes described as the 'negative therapeutic reaction' (Freud 1923). This was so before the distinctive function of K was identified and theorized by Bion. The difficulty in these clinical descriptions of identifying a distinct kind of 'epistemic anxiety' is that the massive anxieties which are shown to be present seemed to be provoked by their particular disavowed objects, not by the experiences of curiosity or learning in themselves. Perhaps the problem is that in these situations, the dimensions of L and H and K are so deeply entwined with one another that it becomes difficult to separate K from its intentional objects, thought from what it is about. It is possible that psychoanalysts may be able to further clarify these differences by reference to their clinical experience.

But even where disturbances in the K function have been the explicit object of attention in Bion's writing, for example in his paper 'Attacks on Linking' (1959), some of these same difficulties, of differentiating between modes of thought and their objects, still arise. The specific 'links' which he saw as being under attack, in the experience of psychotic patients, were the primitive links of the maternal and Oedipal relationships—between breast and mouth, penis and vagina. The disruptions to the capacity for thought which Bion investigated were held to have taken place in the earliest phantasies of the infant, and were then being re-enacted in a primitive kind of transference relationship with the analyst. This came to be understood by him through his close observation of his patients, and through interpretation of the meaning of their fragmented projections.

'Bion', wrote Edna O'Shaughnessy (1981: 60), 'placed the capacity to know (K) at the very centre of mental life. His work puts the pleasure principle and the reality principle on a par with the life instinct and the death instinct as the fundamental governors of mental life'. He believed that L, H, and K could each be the key to the understanding of what is taking place in a psychoanalytic session. Just as the emotions of love and hate take different forms (in the latter case for example, resentment, envy, jealousy, rage, and the refusal to consciously acknowledge them) so there are in Bion's view different kinds of relation to K, or the experience of knowing. He described these as positive, non-existent, or negative: K, no-K, and minus K in his notation. O'Shaughnessy argues, with Bion, that discriminations can and should be made in clinical psychoanalytic sessions, which require deciding how far the central focus of communication lies with the varieties of L, H, or K. Making these distinctions can enable analysts to recognize the central unconscious preoccupations of a patient at that moment, and to be in contact with these. She refers to the anxieties which are evoked by K, which seems to be the equivalent of the idea of epistemic anxiety.

O'Shaughnessy thus long ago described the potential relevance of the idea of K, and thus of anxieties aroused by K, to clinical practice. The idea of 'linking' (Bion 1959) drew attention to attacks on the functioning of the mind, which has particular relevance to the understanding of developmental inhibitions, such as autism and other conditions on the autistic spectrum (Rhode and Klauber 2004), and also to learning difficulties more widely, where these have an emotional or relational dimension. Nevertheless, it seems in practice difficult to discriminate between anxieties about the processes of thought, and those evoked by objects of thought which arouse powerful unconscious feelings. It may not be straightforward to differentiate paranoid–schizoid and depressive anxieties, such as those evoked by Oedipal phantasies, from the anxieties aroused by exposure to knowledge of them. The most primitive mental structures of all may be those in which anxieties about the experiences of knowing, of reality itself, are felt at their most extreme. These include the autistic states explored in Tustin's work, theorized as the 'autistic–contiguous position' in Ogden's (1992), and relating to the psychotic states of mind which Bion (1962) describes. However, individuals who do not suffer from such extreme and chronic conditions may yet have some experience of them during their lives.

Psychoanalytic child psychotherapists have found Bion's work on 'linking' highly illuminating. But again, since they have been seeking to understand states of mind and phantasy being enacted in the consulting room, it is the specific meaning in phantasy of what is being linked, or cannot be linked, which has been given most attention. Thus, a patient who can tolerate the connections of neither words nor numbers, because of the parental coupling which they may unconsciously signify for him, is attacking his K function (his intellectual capability) in a radical way, but this is because of his intolerance of the Oedipal situation. These ideas have been valuable for the understanding of severe disabilities in learning (Simpson and Miller 2004; Sinason 1992) and of the severe aversion to awareness of emotionally significant objects which one finds in autistic states of mind (Tustin 1972). In each of these cases it seems that the unmanageability of the dispositions to love and hate, in the context of the Oedipal situation, is responsible for

disrupting and inhibiting the impulse to learn. Hanna Segal (1957) made the important distinction between the use of symbolism, and symbolic equations, in 'concrete thinking'. An example of a symbolic equation being transformed into symbolic thinking, not from a clinical setting, is the following:

> A young woman working with a small group of emotionally deprived children in an inner city primary school was presenting an account of her work … One very excitable and provocative little boy of six, Patrick, wanted help with his sums. He was one of the children making reasonable progress in reading and was keen to impress his teacher, but one of his difficulties in doing math was that he could never manage to write the figure 8 to his satisfaction. One day, as he struggled to do this in the context of simple addition sums, he began a flow of talk in which he spoke about his teacher's sexual relationship with her husband. In an astonishingly uninhibited monologue, he put into words his excited fantasies about their intercourse. To his teacher's amazement, as his intense preoccupation with intertwining bodies was communicated to her, he suddenly and delightedly drew her attention to the now perfect 8s he could produce. Clearly, the symbolic equation (Segal 1957) in which the joined circles had been equated with joined-up bodies and thus experienced as a source of anxiety had been unpacked, so that separate realms of thinking could exist without confusion

> (Rustin 1999: 128).

As this review demonstrates, the idea which evolved in Klein and Bion's work—that the drive to know and understand is as fundamental an attribute of the mind as the disposition to love and hate—has become influential in contemporary psychoanalysis. Its importance derives in part from developments in the psychoanalysis of children, where experiences of early development and of later severe disturbance through its disruptions of early care have become of pressing clinical interest. And also because of an increased attention, related to relationships both in early life and in psychotic states of mind, in which disturbances of primary mental function are central issues. Clinical phenomena of resistance, denial, splitting, projective identification as a form of evacuation and communication of unbearable emotions and feelings, and dissociation between thought and emotion have needed to be understood. Sometimes this is needed if patients are to be even containable in the clinical setting, and certainly if therapeutic progress is to be made with them. The recognition of the vicissitudes of mental function, somewhat distinct from the emotions of love and hate directed towards their objects in reality and phantasy, has become a valuable theoretical resource for psychanalysts and psychoanalytic psychotherapists. Indeed the ideas of 'containment', 'reflective space', and 'mentalization' (which seeks to enable Bion's ideas about the capacity for thinking of K to become, in a simplified form, a reliably attainable therapeutic goal) have become common terms in psychoanalytic therapeutic practice.[1] The idea of epistemic anxiety—anxiety concerning the felt inadequacy of the mind to situations it faces—arising from the

epistemophilic instinct and Bion's K function, is already implicitly present in clinical thinking and may be capable of further clarification in that context.

Epistemic Anxiety in Society

In the second part of this chapter I shall give attention to the relevance of the concept of epistemic anxiety to wider spheres than that of clinical practice. This rests on the idea that different unconscious anxieties, and social defences against them, are evoked in different institutional settings, and arise from their principal occupational tasks. The idea will be proposed that epistemic anxiety—anxiety about learning—is a kind of anxiety especially liable to be experienced in certain social contexts and situations.

In an edited collection of papers, *Social Defences against Anxiety: Explorations in a Paradigm* (2015), Armstrong and Rustin proposed that just as the anxieties and defences which Menzies had identified in the experience of trainee nurses arose from the distressing nature of their work with patients, so other kinds of work might be expected to give rise to specific kinds of anxiety, and defences constructed to manage these. Several examples of such correlations between work tasks and associated anxieties are developed in that volume, for example in regard to the situations of risk-taking involved in financial trading, and of day nurseries where the distresses of small children have to be contained. Groups whose work involves hard and dangerous physical labour sometimes respond to the anxieties of this situation through creating defensive cultures of endurance and solidarity. There is a parallel to be drawn between the concept of anxieties arising from the tasks of work, and Bion's idea that different basic assumptions are dominant in specific institutions. He referred to the centrality of the basic assumption of 'fight–flight' to military organizations, and the need to manage this in order to enable a rational 'work-group' mentality to prevail among soldiers, and to the role of the basic assumption of dependency in the context of the church. Bion's 'basic assumptions' can be understood as collective defences against anxiety.

The contexts in which 'epistemic anxiety' is most likely to arise and to give rise to defences against it, are educational settings. The idea that curiosity or epistemophilia is itself a natural instinct implies that satisfactions may be gained from its exercise. Epistemophilia means love of knowledge and its concomitant and opposite is hatred of knowledge, what might be termed epistemophobia. These terms correspond to Bion's formulations of K and minus K. Teachers have continuously to enable their students to acquire knowledge and skills, and in doing so to overcome the anxieties associated with learning. Since curiosity and understanding in themselves give pleasure, as well as being potential sources of anxiety, a problem for those involved in educational work is how to ensure that the pleasurable elements of the process prevail over the painful ones. How can it be brought about that epistemic enjoyment outweighs epistemic anxiety?

Freud was aware that the relationship between the 'pleasure principle' and the 'reality principle' is a complicated one. Although the experience of reality itself is often that it frustrates or denies pleasure, the function of the 'reality principle' is a positive one. For

one thing, a grasp of reality may make desires more, not less, capable of satisfaction. And for another, as Klein and Bion understood, the apprehension of meaning can often itself be a source of pleasure. Indeed, lack of opportunity to learn and to satisfy the desires associated with curiosity may itself be a source of pain and distress, giving rise to the state of mind we call boredom, and in its extreme form, depression. Studies of the behavioural systems of mammals, cited for example by Panksepp (1998), have described the significance of the play of young animals in their necessary attempts at learning, for example to hunt. We know from captive and pet animals how lack of stimulus can cause what seems like depression in them. So far as human learners are concerned, there are risks at both ends of this spectrum. If too little is offered to arouse curiosity and interest, there is a stultification of the desire to learn. If what is demanded in terms of learning, or new knowledge, is too much, the learner may be overwhelmed by his task, and turn away from it.

Although curiosity and the desire to learn is an innate human attribute, it functions in definite relations to Bion's other primary dispositions to love and to hate. Knowledge can be pursued in the service of both love and hate. On the one hand, there is the desire to learn about and appreciate the good qualities of 'objects', and on the other, the intrusive and destructive forms of curiosity whose emotional purpose is to diminish and damage them. The epistemophilic instinct which Klein attributed to infants, and its focus on its primary parental objects, was liable to be influenced by loving or hating impulses, at different moments. Curiosity propelled by jealousy involves both a recognition of the admired qualities which the self feels in danger of losing to the possession of others, and antipathy to rivals for the loved one and perhaps to the betraying object itself. The kind of interest in an object which we call envy is specifically dominated by hatred of it. Epistemic anxieties will be shaped to a degree by these interconnections. Destructive curiosity will be liable to arouse unconscious anxieties about retribution and perhaps, if not only hatred is in play, reparative concerns about damage to the object. A form of curiosity or apprehension which is excessively dominated by love for the object may lead to its excessive idealization, and the need to deny aspects of reality which call this in question. The problem in this case is how to avoid a condition of individual or collective narcissism, in which unconscious anxiety that the object may not be perfect may lead to the defence of its further idealization, and the splitting which accompanies this.

Learning or not-learning has a large bearing on everyone's sense of self-worth, from infancy onwards. What children learn and know is an element of their capability and agency in their world. Children's intense awareness of their place in the natural ranking of age, the significance of birthdays, and what 'year' they are in at school, relates to what they believe they are able to do and understand at what stage. Furthermore, their attainments and achievements—from being able to walk and talk onwards—are reflected back in the gaze and behaviour towards them of their objects, be they parents or teachers, which deeply influences how they see themselves. Learning is usually rewarded with approval, or love, and failures to learn risk the withdrawal of approval or love. It is because of the availability of such recognition, approval, and love, as well as from the intrinsic pleasures of learning itself, the satisfactions of the epistemophilic instinct, that

children are persuaded to continually put themselves at risk in new situations of learning, and are prepared to renounce the security of total dependency. Practices of teaching and learning require delicate balances to be maintained between the enhanced satisfactions which arise from learning, and the elements of disapproval and unpleasure which are also deployed to motivate learning. Too much disapproval and dissatisfaction, and learners will withdraw from the field. Undiscriminating approval undermines the reality sense needed for learning to take place. Theories of learning and pedagogy have sought to understand this process. Vygotsky's (1978) 'zone of proximal development'—the idea that learning needs to proceed in comprehensible, sequential stages, so that the next step follows coherently from the last—is one valuable idea. James Strachey's (1934) seminal account of interpretation in psychoanalysis took a related view. He argued that the role of unconscious phantasy in analysands' perceptions can only be revealed to them, and their resistance can only be overcome, through the demonstration of small discrepancies between their perception of the analyst and the manifest reality. He describes here the way in which an analyst can take account of epistemic anxiety as it arises when interpretations are offered. In this instance the resistance is relinquishing a false kind of knowledge—a phantasy—on which at that moment his sense of himself may have depended. Unconscious anxieties involving both love and hate may be involved in such a situation, but it seems that the experience of recognition gives rise to anxiety in itself.

The role of formal education has grown to a large degree in modern industrial and post-industrial societies. Formal education may begin as early as three or four years of age, where governments (as in England) decide that pre-school provision should have qualified teachers among its providers. In many nations, around half of the relevant age group are enrolled in tertiary education, in universities, or other institutions. Occupational opportunities are highly correlated with educational achievements. This situation gives rise to immense social and psychological pressures within educational systems, since the rewards for success and the penalties for failure have become so large.

Michael Young (1958) described, in *The Rise of the Meritocracy*, a dystopian future in which domination would be exercised by those who had proved their merits through their educational success, while those in lower positions were left to blame themselves for their failures in a system which claimed to provide opportunities for all. His suggestion was that in some ways inequalities which arise from competitive systems based on competitive achievement could be more psychologically painful to the unsuccessful than those which seem to have occurred merely through accidents of birth. A meritocratic apportionment of success and failure seems similar in some respects to a Calvinistic distinction between the elect and the damned, the elected in meritocracies being held to have demonstrated their 'chosen' status through the educational certificates they acquire. (Of course, we know that ostensibly meritocratic social systems are often deeply compromised by the advantages which the offspring of some social groups, independent of their individual talents, are able to bring to educational competition).

Epistemic Anxiety

Different strategies are adopted by societies and institutions in response to the 'epistemic anxieties' inseparable from contexts of learning. Highly competitive systems may divide not only the rewards of learning, but also the penalties and anxieties arising from failure to learn, between one group and another. Systems of ranking and grading have this divisive effect, since those at the top gain assurance of their capability and position of honour, while those placed lower down are left to bear the anxiety and shame of being unable to learn to the required standard. In the English educational system, schools and universities as well as pupils are subjected to systems of grading and ranking (through inspections, 'league tables', and the like), which has the effect that competitive pressures and their anxieties are projected by institutions into their students. We know from research into educational selection and streaming (Jackson 1964) that the mere designation or 'labelling' of children as less capable itself causes a lessening of their capability. More inclusive approaches, more influential in the primary than the secondary sector of British education, seek to support all children's ability to learn. Critics of such systems, especially in secondary education where 'comprehensive' schools at one time embodied such a conception, argue that unless discriminations are made between capacities to learn, and priority accorded to the most capable, high levels of achievement by the most capable students will be diminished. According to this view, 'standards' will then suffer, although whether these are the standards of the most capable, or of the majority, or even of the least capable, is rarely specified. Much public debate about educational policy and practice reflects these different priorities.

A distinction can be made between the anxiety and shame that can be evoked by the experiences and difficulties of learning itself, and that which can be evoked by encounters with specific areas of knowledge, for example where these are fields of contention. For example, attempts to reflect on issues of race, sexuality, and gender in educational settings, and elsewhere, may give rise to specific epistemic anxieties, as individuals and groups may fear being shamed for holding unacceptable views or for their supposed ignorance.

While epistemic anxiety has its most visible presence in educational settings, it seems that in some contemporary societies it has become a phenomenon of political significance. Whereas social antagonisms in an early period were mainly structured by differences of income and wealth, it appears that today these have been reordered to reflect differences of educational advantage. The most significant demographic difference between those who voted for Trump and those who voted for Clinton in the 2017 United States presidential election lay in the proportions of those with and without college education (Edsall 2017). In part, this reflects the fact that power in modern societies is largely exercised by those with high levels of formal education. It may be that those perceived to be privileged and neglectful of the interests of the middle or working class were resented and opposed because of their relative power and wealth, not because of their educational achievement. But this seems not to be entirely the case. What happened does seem in part to have been a revolt of the less against the more educated. Trump's election campaign involved deliberate attacks on the idea that political debate should be conducted according to the 'codes' which attribute value to consistency and the proficient

deployment of evidence in arguments. These codes are subscribed to by 'quality' media, the law, the scientific establishment, and have been established conventions of political debate. This entire way of thinking, Trump in effect asserted, is a source of oppression for the 'ordinary' (less educated) American. Attacks on the autonomy of the courts, on climate science, and on leading newspapers and television networks for their alleged biases and disregard for popular opinion, have been elements of this assault. Trump's choice of 140-character 'tweets' as his preferred mode of communication with publics surely reflects his disdain for more 'difficult' and elaborated forms of thought and expression. In Bion's terms, one could say that they have the function more of the evacuation of angry feelings than communication. Similar phenomena were observed in the European Referendum campaign in Britain in 2016, in assertions that too much attention had been customarily paid to 'experts' and 'evidence', with the implication that collectively shared feelings might be a more reliable guide to the truth and to proper action. They seem to be part of the rise of populism and nationalism more widely.

These campaigns surely gave voice to a feeling of injury and insult on the part of populations who felt that their own less sophisticated and 'authorized' forms of thought and expression had disqualified them from being listened to by conventional politicians and authorities. We can understand what is happening as a vehement reaction to the perceived rise of a privileged meritocracy. Richard Sennett and Jonathan Cobb (1972) once described the 'Hidden Injuries of Class'; we are now able to observe the effects of the hidden injuries of education, to those who are excluded by it.

There is a long tradition of thought which has regarded 'mass society', and indeed the emergence of democracy, with its media of mass communication, with suspicion and foreboding. Theodor Adorno's (1951) essay on fascist propaganda was a classic demonstration of how popular sentiments could be manipulated through the mass media of the press and broadcasting. Media of communication consistently offer to their audiences opportunities both to satisfy impulses to learn, and to deflect, displace, and resist these, as much through offered varieties of trivia and repetition as in the deliberate negation and subversion of rational forms of understanding. In so far as learning is usually accompanied by anxiety, it is to be expected that some of the time nearly all individuals will seek refuge from its stresses in kinds of mental function which demand little of them. These two contrasting forms of communication—those committed to learning, those averse to it—are in a constant alternation with one another in the outputs of the broadcasters, the press, and the entertainments industry. It remains unclear whether the greater opportunities for communication between individuals and groups which are provided by the proliferation of its electronic resources—social media, search engines, 'platforms', and 'apps' of innumerable kinds—will make possible an inclusive educated democracy of a kind which has been hitherto unachievable (Wikipedia is an example of where widely felt desires for greater knowledge and understanding are being met), or will give rise instead to varieties of 'soft totalitarianism', in which most communication is covertly controlled by elites.

Conclusion

The aim of this chapter has been to bring together the psychoanalytic theories of anxiety and of epistemophilia, to see if their conjugation would identify a distinctive variety of anxiety associated with learning. Insofar as this is constituted by the concept of 'epistemic anxiety' an additional question arises. Can these conjoined theories enable us to understand how epistemic anxiety can most effectively be managed, in order to enable learning to be facilitated rather than undermined and resisted? The answer to this question, at every level of the learning experience, from infancy to adult political life, may lie in Bion and his successors' theory of containment, or container–contained relations. Crucial to the experience of learning, in all its phases and locations, is the recognition of the anxieties which accompany it, of which one specific kind is the capacity to learn itself. It is where this anxiety is recognized and understood that learning will best take place.

Epistemic anxiety is one of several forms of anxiety which have relevance to the spheres of social and political philosophy. One can differentiate and choose between societies and social institutions in regard to the way in which they 'organize' anxieties and their productive and destructive functions. On the positive side lie the motivating role of anxieties in various spheres of competition and enterprise. On the negative side, there are the states of affairs where anxieties become excessive and give rise to the defences of social phobias and paranoias. Such an approach to anxieties brings emotions, both collective and individual, into the centre of social and political thinking, and thereby identifies a key space for psychoanalysis as a way of understanding their unconscious presence and force.

References

Adorno, T. (1951). 'Freudian Theory and the Pattern of Fascist Propaganda'. Reprinted in J. Prager and M. Rustin (eds), *Psychoanalytic Sociology*, Vol. 1 (pp. 115–134). London: Edward Elgar.

Alford, C. F. (1997). *What Evil Means to Us*. Ithaca, NY: Cornell University Press.

Armstrong, D. (2007). *Organisation in the Mind*. London: Routledge.

Armstrong, D. and Rustin M. J. (2015). *Social Defences against Anxiety: Explorations in a Paradigm*. London: Karnac.

Bick, E. (1968). 'The Experience of the Skin in Early Object-Relations', *International Journal of Psycho-Analysis 49* (1): 484–486.

Bion, W. R. (1959). 'Attacks on Linking', *International Journal of Psychoanalysis 40*: 308–315.

Bion, W. R. (1961). *Experiences in Groups*. London: Tavistock Publications.

Bion, W. R. (1962). *Learning from Experience*. London: Heinemann.

Bion, W. R. (1970). *Attention and Interpretation*. London: Heinemann.

Bléandonu, G. (1994). *Wilfred Bion: His Life and Works 1897–1979*. London: Free Association Books.

Britton, R. (1993). 'Fundamentalism and Idolatry'. In C. Covington, P. Williams, J. Arundale, and J. Knox (eds), *Terrorism and War: Unconscious Dimensions of Political Violence*. London: Karnac.

Britton, R. (1998). *Belief and Imagination*. London: Routledge.

Britton, R. (2015). *Between Mind and Brain: Models of the Mind and Models in the Mind*. London: Karnac.

Edsall, T. B. (2017). '**The End of the Left and the Right as We Knew Them**'. *New York Times* 22 June 2017.

Breuer, J. and Freud, S. (1895). *Studies on Hysteria*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 2. London: Hogarth Press.

Freud, S. (1905). *On Psychotherapy*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol 7 (pp. 257–268). London: Hogarth Press.

Freud, S. (1909a). *Analysis of a Phobia in a Five-Year Old Boy ('Little Hans')*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 10 (pp. 3–149). London: Hogarth Press.

Freud, S. (1909b). A *Case of Obsessional Neurosis ('the Rat Man')*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 10 (pp. 153–320). London: Hogarth Press.

Freud, S. (1917a). *Mourning and Melancholia*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 237–258). London: Hogarth Press.

Freud, S. (1917b). *Introductory Lectures on Psychoanalysis, Part III: General Theory of the Neuroses*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 16. London: Hogarth Press.

Freud, S. (1923). *The Ego and the Id*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 19 (pp. 3–66). London: Hogarth Press.

Freud, S. (1926). *Inhibitions, Symptoms and Anxiety*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 20 (pp. 74–176). London: Hogarth Press.

Halton, W. (2015). 'Obsessional-punitive defences in care-systems: Menzies Lyth revisited'. In D. Armstrong and M. J. Rustin (eds), *Social Defences Against Anxiety: Explorations in a Paradigm* (pp 27–38). London: Routledge.

Henry, G. (1974). 'Doubly Deprived'. *Journal of Child Psychotherapy 3*(4): 15–28.

Jackson, B. (1964). *Streaming: An Educational System in Miniature*. London: Routledge & Kegan Paul.

Jaques, E. (1951). *The Changing Culture of a Factory: A Study of Authority and Participation in an Industrial Setting*. London: Tavistock Publications (repr. London: Routledge, 2001).

Jaques, E. (1955). 'Social systems as a defence against persecutory and depressive anxiety'. In M. Klein, P. Hermann, and R. E. Money-Kyrle (eds), *New Directions in Psychoanalysis* (pp. 478–498). London: Tavistock Publications (repr. London: Karnac, 1985).

Khaleelee, O. (2003). '**Not leading followers, not following leaders; the contemporary erosion of the traditional social contract**'. AK Rice Institute.

Khaleelee, O. and Miller, E. J. (1985). 'Beyond the small group: Society as an intelligible field of study'. In M Pines (ed.), *Bion and Group Psychotherapy* (pp. 354–385). London: Routledge & Kegan Paul.

Klein, M. (1927). 'Criminal Tendencies in Normal Children', *British Journal of Medical Psychology 7*: 177–192. (Reprinted in Klein 1997a).

Klein, M. (1928). 'Early Stages of the Oedipus Complex', *International Journal of Psychoanalysis 9*: 167–180. (Reprinted in Klein 1997a).

Klein, M. (1930). 'The Importance of Symbol Formation in the Development of the Ego', *International Journal of Psychoanalysis 11*: 24–39. (Reprinted in Klein 1997a).

Klein, M. (1931). 'A Contribution to the Theory of Intellectual Inhibition'. *International Journal of Psychoanalysis 12*: 206–218. (Reprinted in Klein 1997a).

Klein, M. (1934). 'On Criminality'. *British Journal of Medical Psychology 14*: 312–315. (Reprinted in Klein 1997a).

Klein, M. (1935). 'A Contribution to the psychogenesis of manic-depressive states', *International Journal of Psychoanalysis 16*: 145–174. (Reprinted in M. Klein 1997a).

Klein, M. (1940). 'Mourning and its Relation to Manic-Depressive States', *International Journal of Psychoanalysis 21*: 125–153. (Reprinted in M. Klein 1997a).

Klein, M. (1945) 'The Oedipus Complex in the Light of Early Anxieties', *International Journal of Psychoanalysis 26*: 11–33. (Reprinted in Klein 1997a).

Klein, M. (1946). 'Notes on some Schizoid Mechanisms', *International Journal of Psychoanalysis 27*: 99–110. (Reprinted in M. Klein 1997b).

Klein, M. (1997a). *Love, Guilt and Reparation and Other Works 1921–1945*. London: Vintage.

Klein, M. (1997b). *Envy and Gratitude and Other Works 1946–1963*. London: Vintage.

Meltzer, D., Bremner, J., Hoxter, S., Weddell, D., and Wittenberg, I. (1975). *Explorations in Autism*. Perthshire: Clunie Press.

Meltzer, D. and Harris Williams, M. (1988). *The Apprehension of Beauty*. Perthshire: Clunie Press.

Menzies, I. (1960). 'A case study in the functioning of social systems as a defence against anxiety: A report on a study of the nursing service of a general hospital'. *Human Relations 13*(2): 95–121. (Reprinted in Menzies Lyth 1988).

Menzies Lyth, I. (1988). *Containing Anxieties in Institutions: Selected Essays*. London: Free Assocation Books.

Ogden, T. (1992). *The Primitive Edge of Experience*. London: Karnac.

O'Shaughnessy, E. (1981). 'A Commemorative Essay on W. R. Bion's Theory of Thinking'. *Journal of Child Psychotherapy 7*(2).

Panksepp, J. (1998). *Affective Neuroscience: The Foundations of Human and Animal Emotions*. Oxford: Oxford University Press.

Rhode, M. and Klauber, T. (2004). *The Many Faces of Asperger's Syndrome*. London: Karnac.

Rosenfeld, H. (1987). 'Destructive Narcissism and the Death Instinct'. In *Impasse and Interpretation* (pp. 105–132). London: Tavistock.

Rustin, M. E. (1999). 'The Training of Child Psychotherapists at the Tavistock Clinic: Philosophy and Practice'. *Psychoanalytic Inquiry 19*(2): 125–141.

Rustin, M. E. and Rustin, M. J. (2017). *Reading Klein*. London: Routledge/Institute of Psychoanalysis.

Rustin, M. J. (2009). 'How do psychoanalysts know what they know?'. In L. Braddock and M. Lacewing (eds), *The Academic Face of Psychoanalysis*. London: Routledge.

Rustin, M. J. (2015). 'Anxiety and Defences, Normal and Abnormal'. *Organisational and Social Dynamics 15*(2): 233–247.

Segal, H. (1957). 'Notes on symbol formation'. *International Journal of Psychoanalysis 38*: 391–397.

Sennett, R. and Cobb, J. (1972). *The Hidden Injuries of Class*. New York: Alfred Knopf.

Simpson, D. and Miller, L. (eds) (2004). *Unexpected Gains: Psychotherapy with People with Learning Difficulties*. London: Karnac.

Sinason, V. (1992). *Mental Handicap and the Human Condition*. London: Free Association Books.

Steiner, J. (1993). *Psychic Retreats: Pathological Organisations in Psychotic, Neurotic and Borderline Patients*. London: Routledge.

Strachey, J. (1934). 'The Nature of the Therapeutic Action of Psychoanalysis'. *International Journal of Psychoanalysis 15*: 127–159.

Tustin, F. (1972). *Autism and Childhood Psychosis*. London: Hogarth Press.

Vygotsky, L. (1978). '**Interaction between learning and development**'. In L. Vygotsky, *Mind and Society* (pp. 79–91). Cambridge, MA: Harvard University Press.

Young, M. (1958). *The Rise of the Meritocracy*. London: Thames and Hudson.

Notes:

(1.) [Eds: For further discussion of mentalization and the development of thought, though in the context of literature, see Galgut, this volume.]

Michael Rustin

School of Law and Social Sciences, University of East London

Oxford Handbooks Online

Abstract and Keywords

This chapter is concerned with exploring the relevance of gender as a critical category for clinical psychoanalysis. It recognizes two developments in the late twentieth century that may appear to present gender as either not a pressing issue in self-development or as sufficiently troubled and undone to have lost its regulatory grip. The first concerns the domination of the psychoanalytic imagination with preoccupations other than sexuality, sexual difference, and gender; and the second is linked to the deconstruction and reconstruction of hetero-normative gendered frameworks initiated by cultural gender theorists. It is argued that the gendered binary of Western thought with its socially normative values and assumptions shapes the unconscious minds of every person. Notwithstanding critical appreciation of the gendered discourses of psychoanalysis as well as expanded thinking about the possible repertoire of individual gender variations, gender continues to carry evaluative burdens.

Keywords: gender, sexual difference, gendered discourses, clinical psychoanalysis, power

Louise Gyler

(p. 709) Introduction: Psychoanalysis, Gender, and Power

GENDER issues impact individuals in profound, indefinable, and contradictory ways, both in public and private life, especially as power dynamics are so intimately involved in the production of gendered meanings. Historically, psychoanalysis has been prone to getting it wrong when it comes to gender matters—for example, women, same-sex desire, and trans experience—by overlooking the workings of power in the production of knowledge and gender inequalities as they find expression in psychic life, clinical practice, and

theorizing. Social and cultural norms have been conflated with psychoanalytic theory; description with prescription. The question that I am asking with 'Does gender matter?' is whether psychoanalysis is now suitably cognizant of the workings of gender binary logic of Western thought so that the consequences of the more controversial aspects of our past can be grappled with psychoanalytically; are we (psychoanalysts) able to submit our theoretical and personal explicit and implicit biases to analysis so that enough of any non-reflexive tendencies are ameliorated? Have we arrived somewhere approximating the 'third attitude' advocated by Julia Kristeva (1979/1986) in 'Women's Time' all those years ago? For Kristeva, the first phase is associated with sociopolitical demands of women for equality and the second with 'the specificity of female psychology and its symbolic realisation' (1986: 194). Kristeva wrote:

> In the third attitude, which I strongly advocate—which I imagine?—the very dichotomy man/woman as an opposition between two rival entities may be (p. 710) understood as belonging to *metaphysics.* What can 'identity', even 'sexual identity', mean in a new theoretical and scientific space where the very notion of identity is challenged? (1986: 209)

Kristeva goes on to argue that she is not proposing a hypothetical bisexuality as that would only serve to erase difference but rather suggesting 'the demassification of the problematic of *difference'* (1986: 209). Virginia Goldner seems to think that the theoretical grip of sexual/gender difference has weakened: 'Nearly four decades of feminist and queer theory have disabled the sex/gender binary, undermining the assumption that something called "gender" imposes its cultural will on a pre-existing, universalized "sexed body" ' (2011: 160). The above propositions lead to questions about where we are now: are we in Kristeva's imaginative place where the distinction between women and men no longer carries the normative evaluative freight that it historically has; have we arrived at a place, at least theoretically, where gender and the biological body are no longer naturalized as in unison as Goldner claims?

Recently when discussing with psychoanalyst colleagues questions concerning gender biases in our theory and their implications for practice, I was met with puzzlement as my colleagues appeared to feel that these issues were no longer of relevance for contemporary practice, believing we had worked through the problems and prejudices of previous decades. This led me to ask: has gender been sufficiently 'troubled' and 'undone' to no longer have purchase in (some) psychoanalysts' consulting rooms? Has gender been sufficiently disambiguated from the effects of the workings of power relationships? Has the regulatory gaze with its attendant anxieties for both psychoanalyst and analysand been sufficiently relaxed? On one level, I am not unsympathetic to my colleagues' position because not only have our attitudes to gender and sexuality been reworked but there have been paradigm shifts in psychoanalytic thinking (and here I refer primarily to Anglo-American trajectories). These paradigm shifts, in one sense, potentially leave sexuality and gender behind, as André Green lamented in his provocatively entitled paper 'Does sexuality have anything to do with psychoanalysis?' (1995), and position other clinical preoccupations as more central and urgent. I shall suggest that, with these

developments, although the meaning of gender and gender discourses in the consulting room is both hard to pin down and writ large, gender theory continues as critically relevant for psychoanalytic practice.

Questions of gender function on various interrelated levels in clinical psychoanalysis—these include the level of individual psychology; the theoretical level; and the way individual psyches, theoretical ideas, and differing styles of participation in the consulting room interweave to produce gendered psychoanalytic discourses (Gyler 2010; Mitchell 1996). To explore the significance of gender and gender theory for contemporary practice, I briefly touch upon issues such as: where have we been on questions on gender? where are we now? and where might we be going? These questions involve the complicated intersections between psychoanalysis and cultural gender theory. Before discussing these issues, I propose to sketch briefly the contemporary landscape of clinical psychoanalysis to give some context for my colleagues' perplexed attitudes.

(p. 711) **Contemporary Clinical Concerns**

Psychoanalytic developments since the second half of the twentieth century have been marked by preoccupations other than sexuality and gender; notably object relations, attachment and loss, and the development of thought have dominated the psychoanalytical imagination. I suggest that contemporary clinical psychoanalysis, no longer deploying an 'archaeological' model or even particularly concerned with reconstruction (of the primal scene), aims at the work of representation—as met with, for example, in Bion's theory of thinking (1959, 1962a, 1962b), Botella and Botella on the work of psychic figurability (2005), and Roussillon (2011) amongst others. Bion's theory of thinking is based on four overlapping and interconnected principles of mental functioning. They are: a person's need to know the 'truth'; that two minds are required to think a person's disturbing thoughts (this principle implicates the mechanism of projective identification as communicative); a capacity for thinking as necessary to manage constructively the thoughts and feelings generated from disturbing experiences (this principle relates to the model of container/contained); and, 'dreaming' conceived as a psychoanalytic function of the personality that occurs in wake and sleep and is necessary to work through emotional issues. More recently there is a developing literature based on Bion's later work (1970, 1992) on the unrepresented and unknowable (called O) that represents the limits of knowledge as acquired through the senses and can only be experienced in the becoming at-one (Eshel 2017). Put simply, psychoanalysis is concerned with the problem of how to think and with experiencing deeper levels of psychic reality rather than with the question of what to think (that is, making the unconscious conscious which assumes the existence of formed but defended thought). Of course, all levels of psychic life fold together in an analysis but I am suggesting that contemporary attention rests with the problems of representation and psychic states of being-in-oneness. Clinically, this means that there is increased attention in the literature to presymbolic manifestations, for example children who are unable to play (Tustin 1988), autistic transformations (Korbivcher 2014), and enactments of unverbalized/

unrepresented experience. Careful consideration is given to meaning-making in the 'here' and 'now' intersubjective encounter (Joseph 2013; O'Shaughnessy 2013). These modified clinical aims appear to position gender and sexuality as no longer 'bedrock' in the currency of the consulting room.

Along with the interest in archaic states of mind and with analytic engagement with patients who function in the presymbolic and non-verbal realm of experience, there has been a shift in the analyst's mode of participation in the analytic encounter. As a generalization, it is less authoritative, relying less on universalizing forms of knowledge. In practice, careful consideration is given to how explicit and implicit theoretical commitments and personal responses, attitudes, and biases all form part of how we emotionally receive and make sense of our analysand's communications—and the capacity for emotional resonance is needed when it comes to shaping responses to unrepresented experience. In addition, there has been a changed attitude to theorizing (p. 712) itself marked by a tendency to shift away from universalizing accounts especially of gender and sexual development to constructing knowledge on more particular and individualized bases as Nancy Chodorow promotes:

> We need to move beyond those theories that base their psychologies of gender on a search for gender differences and a comparison of men and women or that reduce gender psychology to one element (whether genital, relational, or anything else). We need a theory that can point us toward and help us understand the clinical individuality of gender.

> (Chodorow 2012: 43)

The Intersubjective Space: Clinical Examples

The following clinical vignettes are illustrations of how the issue of gender and sexuality may not appear pressing in the lives of some children or in the foreground of the analytic working through in the consulting room.

Mandy, aged ten years, began her analytic sessions by creating with the play materials an imaginative island world inhabited with various 'creatures' who would cook up exotic meals, sleep, play, fight, work, travel, and so on. For several months at the beginning of her analysis, Mandy rarely spoke. Whenever I spoke to explore and/or link the meaning of her play with emotional states and/or what might be happening in the room between us, she appeared to ignore any of my overtures, remaining silent. After a period, her play and our interaction began to feel repetitive and the emotional atmosphere in the consulting room began to alter. I sensed her increasing anger, which appeared to resonate with my own experience of frustration, confusion, and sense of failure at being unable to make any emotional contact with Mandy. I guess we both suffered for a time before I could interpret (with emotional conviction) that she heard my words as confusing and foreign, leaving her feeling angry with me for intruding into her world with my bewildering words

and ideas. This interpretation about the 'here' and 'now' of our interaction did connect with Mandy who was able to acknowledge her agreement and, from this juncture, in her analysis, she began to talk freely to me. Much later, she could confide that this moment in her analysis marked an important milestone for her; she could have a conversation with another person and that was a great relief to her.

What stops conversations or dialogue—an all too common feature of contemporary analyses? Why didn't Mandy tell me that my words, or my presence for that matter, were confusing to her so we could have explored this together—a coming together rather than a turning away? On a manifest level, we might consider that she simply did not trust me to listen to her experience—and, yet, she steadfastly attended her sessions, even when occasionally they clashed with apparently more appealing activities. In the transference, I embodied a confusing and confused object/figure that anxiously needed to intrude into her, disturbing her equilibrium, and frightening her. Mandy defended (p. 713) against this intrusive object by shutting down, with this dynamic externalized and enacted in the intersubjective encounter. As this non-verbal communication came to be understood and contained, verbal representation and speech became possible. In Mandy's analysis, the transference/countertransference experience needed to be lived in the consulting room and experienced as emotionally 'real', before the analyst could speak cogently about the experience to her.

To give another example: I have described elsewhere the experience of seeing a small boy, James, aged two years, who had been exposed to serious violence in his primary caregiving relationship, which he then enacted with a series of multiple caregivers (Gyler 2010). Not unexpectedly, in his sessions he related to me with fear and actual aggression. In one session, after I had registered both in my mind and in my physical reaction (I pull back my hand to protect it) my alarm about being physically hurt (scratched by toy scissors), I was able to speak to him about his terror at being with me. He then responded: 'Who are you?' This locution allowed a space for clarification and interpretation after which the small boy could reveal playfully his fear of his falling apart 'world'. The question, 'who are you?' is a question about identity, about in this instance the identity of the other—is she who I think she is? Is she a dangerous murderous (m)other? Is she different from my own 'fixed' phantasmal ideas of others? Who then is 'she' and who am 'I'? These are questions of identity that this small child needed to explore. The issue of whether there are or can be satisfactory and satisfying answers, or whether the 'answers' are provisional and shift over time and place, is another matter. Nonetheless, exploring these questions may involve anger, disappointment, and mourning and a confrontation with the limits and constraints of human freedom.

In these clinical vignettes, the children begin to think and articulate their wishes and fears rather than enact them in unproductive and/or destructive forms of relatedness: the wish to speak in a silent child and the wish to be curious about the other in a child who attacked others. These wishes are relational but not gender related in any direct sense. These children require the recognition of the other so that they can develop their capacity for thought. With the capacity for symbolic thought comes the painful

confrontation of needing to tolerate and mourn the inner reality of their damaged internal object. I have argued elsewhere (Gyler 2010) that, when difficulties associated with the primary caregiving relationship are in relation to women as mothers, this can result in negative and frightening internal representations of the maternal object if these types of representations are not mourned and balanced with positive ones. This internal psychic situation can lead, via the symbolic associative chain of maternal/mother/women, to deprecating attitudes towards women. In this sense, these problematics implicate a gendered discourse. For the individual, gendered meanings are produced by and productive of experience where the psyche and the social interact 'enigmatically' to fix, resist, and refuse the normative gender binary.[1]   (p. 714) Psychoanalysis as a critical self-reflexive endeavour holds the potential to analyse such ongoing effects produced by gendered discourses.

Gender Discourses: Past into Present

Until the mid-twentieth century, gender was mostly used to refer to grammatical categories. The genesis of gender as analytical concept, perhaps, begins with Simone de Beauvoir's enquiry: why is Woman Other? (1949/1972). This incisive question cuts to the heart of the problem of women's otherness and their disadvantaged location in the binary system of Western thought. The term 'gender' was introduced into the medical field by John Money (1973) when he made the distinction between sex (biologically given) and gender (socially constructed) in the 1950s while investigating sex roles and gender-related behaviour of individuals with intersex conditions. Robert Stoller (1968) introduced the term into psychoanalysis through his work on sexual/gender identities. The widespread uptake of the term, however, is associated with second-wave feminism where in the UK, Ann Oakley (1972) and in the USA, Gayle Rubin (1975) and Rhoda Unger (1979) initiated the use of the sex/gender distinction for feminism.

Freud can be read as not interested in gender identity as we would understand it today because for him, gender and sexuality were not bifurcated as they mostly are today. Nevertheless, psychoanalytic usages of the term 'gender' are closely associated with the contested space of sexual difference. Freud's account of sexual difference, the question of how desire and identifications are structured so lives are lived as sexually differentiated, clearly has implications for our contemporary gender thinking. Freud (1905) offered a finely textured but unstable account of sexuality with the biological body and sexual desire being motivationally fundamental to the development of mind.[2] His writings hold contradictory claims so that subsequent authors can read off various passages to support both socially conventional and more radical positions in relation to gendered functioning. For example, his radical insights include that there are no simple corresponding relations between libidinal instincts and their aim and objects and that 'pure masculinity and femininity remain theoretical constructs of uncertain content' (Freud 1925: 258). As we know, there are also normative claims of what it means to be a man or woman (Freud

1925, 1931, 1933). For Freud, sexual differentiation with its binary logic is central for the construction of subjectivity and entry into the symbolic order.

In the nineteenth century with the advent of the modern medical interventions and the shift from theological moral views of mental distress to medicalized pathological views, 'bad' women were reframed 'mad' women (Appignanesi and Forrester 1993; Showalter 1985). These developments meant that, for example, women who resisted and (p. 715) refused normative gender roles could be 'medicalized' and, for example, have their reproductive and sexual organs removed as cure for their apparent derangements. In this context, 'the talking cure' introduced a fundamentally different approach, permitting a space for women to voice their sexual desire. Nevertheless, Freud's account of the possible outcomes for women's development was negatively freighted—for example, they are more masochistic, narcissistic, jealous, envious, with weaker and less independent superegos than men, and their 'nature is determined by their sexual function' (Freud 1933: 135). In her landmark text, Juliet Mitchell (1974) introduced the argument that Freud was describing the limiting conditions for women under patriarchy rather than prescribing how women should be. The description/prescription debate continues (Mitchell 2015). From the outset, Freud's accounts of the psychology of women was contested and analysts such as Jones (1927), Klein (1928), and Horney (1925) shed light on the 'dark continent' (Freud 1926: 212) in Freud's lifetime. Subsequently, other writers have theorized a representational space for women's ambition, aggression, and agency (Elise 2008; Harris 1997), to include representations of the female body (Balsam 2012) and the significance of the representation of prediscursive phenomena for subjectivity, especially inner spaces to hold the possibility of (pro)creativity being called into attention (Gyler 2017).

Over the past century or so, termed the 'sexual century' by Ethel Person (1999) and perhaps book-ended between *The Three Essays* (Freud 1905) and *Gender Trouble* (Butler 1990), classical psychoanalytic insights concerned with sexuality and gender have been challenged, deconstructed, and reconstructed to render them more available for the representation of human freedom. We have seen psychoanalytical critical interventions into the woman 'question' (described earlier), theorizing same-sex desire, gender binary discourses, and ways of conceiving queer and trans experience. Accounts of same-sex desire aimed at depathologizing homosexuality have been productive along with the theoretical move to problematize all sexualities (Chodorow 1994). Ideas such as 'gender as destiny' as the bedrock of psychoanalytical discourse (Young-Bruehl 1996) and notions of a binary construction of gender have been challenged with gender being reconceived as relational, provisional, in-process, contingent, assembled, fluid, and indeterminate (Benjamin 1996; Harris 2005). The need for a representational space for trans and queer experience and the delinking of any naturalized connection between gender identity and the anatomical or natal body have been thoughtfully discussed (Goldner 2011). The role and status of the body remains uncertain and contentious in psychoanalytical discourse but interventions that, for example, conceive bodies as 'hybrids of natures-cultures' (Flax 2006: 1134) and present a symbiological approach where nature/nurture and biology/culture inextricably determine one and another (Gagnier 2017) are well established. In

these shifts, not only is the universal heteronormative male subject of Western thought more precariously seated in his occupation of the privileged epistemological and conceptual position, but also conceptions of 'gender as destiny' and sexual difference are in contention; for example, we can now ask if we need to be a sexually differentiated subject to be able to function symbolically.

(p. 716) In contemporary psychoanalysis and cultural discourses, sexual difference is understood within two broad frameworks as Scott (2011) proposes. In the first, sexual difference is seen as a given, that is, a natural bodily based phenomenon, on which the gender system is culturally constructed (for example, Chodorow 1978). In the second, sexual difference is viewed as the effect of historically variable discursive practices of gender which produce the body as sexual. This position has been most notably advocated by Butler (1990). Following Freud and Lacanian readings of Freud, sexual difference is given meaning through fantasy and identification to produce either masculine or feminine psychic identities. Anatomical bodies are seen as not necessarily corresponding with the feminine and masculine psychic positions, that is, these positions are not viewed as directly expressing any natural biological distinction. The categories of feminine/masculine and in association female/male are, however, asymmetric in their relationship to the signifying processes of subjectivity. They have a different relationship to the phallus, with the feminine position standing as the negative term—absence/castrated, and the masculine for the positive term—presence/phallic; that is they accrue and restate meaning from the gendered binary of Western thought. For example, one of the effects of this discourse can be seen in the literature on 'the paternal function', a function conceptualized to intervene to promote psychic separation in the primary maternal dyadic relationship. The paternal function becomes associated for example with knowledge (Perelberg 2013) and time (Aisenstein 2015), perpetuating a gendered theoretical discourse whereby the fantasy of the masculine/male's association with culture and reason and the feminine/female's association with nature and emotion is sustained. While not disputing the role and function of the third-party function and the need for difference, other writers have questioned the use of such gender-specific language as it risks conflating phenomenological experience and fantasy with theoretical abstraction (Glocer Fiorini 2010, 2017; Gyler 2010). I suggest that this gender-specific linguistic style is problematic as it appears to claim that the patriarchal version of the symbolic order is immutable and not open to contestation rather than one more contingent and entwined with social practice as Judith Butler argues (2004).

Along with developments in rethinking gender and sexual difference also comes uncertainty. We start to believe that we understand our terms, categories, and references, only to find the definitions and understandings have been modified and we are in a constant state of anxiety that we will never capture with sufficient acuity, subtlety, and sensitivity the multitude of nuances we need to communicate about gender, sex, the body, sexuality, and sexual difference. 'Gender' now is an out-of-focus term situated in a web of other terms such as woman, man, femininity, masculinity, sex, sexuality, body, embodiment, sexual difference, the social, and culture. It refers to an analytical, political, sociological, and psychological concept. Although contemporary gender theory has

displaced or even replaced the category of woman as central, I argue that the problem of women and gender continues as a substantive issue—notwithstanding the conceptual expositions of gender from feminist, queer, and trans theorists in recent decades (see Corbett 2008 for an overview of these developments) and attempts to generate 'new' languages for our ever-expanding ways of living and 'doing' gender.

(p. 717) Gendered Desires and Dis-ease: A Clinical Example

I now turn to clinical material in which desires and anxieties in relation to gender and the body are in the foreground of the clinical presentation and the work in the consulting room. Sarah, in her twenties, sought analysis because she was troubled by separation anxieties and her worries about the meaning of her subjectively experienced bodily reaction (she equated her vagina and its secretions with a penis because she believed in an out-of-awareness way that vaginal secretions were active, and activity was associated with masculinity and men and passivity with femininity and women). She held split contradictory self-representations: on the one hand, successful, competent, and superior —'cocky' to use her term; and on the other, chronically anxious and vulnerable and with life-long separation anxieties. Not unexpectedly, from early in her analysis, she responded to interpretations that reflected on the analyst's difference and the asymmetry of the analytic relationship, as the analyst being more interested in analytic technique than in her. She made clear her wish for a seamless relationship where everything could be arranged on her terms (e.g. fees, appointment times, and the ending of sessions). For example, she believed that the analyst had a difficult job because the analyst was required to end the session without patients feeling cut off. When no seamless options were available, it signalled her own or her object's failure, leaving her feeling defeated and depressed. She believed that it was impossible to manage two things at once (work/ analysis, etc.). Her wish, however, for an enclosed state with her object endlessly broke down because it was also a terrifying state of mind where she feared losing herself.

Sarah dreamed recurrently that her body possessed the anatomical characteristics of both genders which meant for her that she could psychically position herself as 'complete'. For example, she introduced a dream fragment in which *women not only have vaginas but also little penises*. She elaborated that *the penises are baby ones that do not work properly.* Her association was to a dummy (mannequin) in a shop display. In the working through of the dream material she became aware of her wish to have everything but also that her belief of having everything was pretence. In another dream, she reported that *she has a penis coming out of her neck and that there is something wrong about it*. In another dream, *she goes to the toilet and does not need to squat as there is a direct flow, which is exciting because it is like a direct flow from a penis.* Her associations to the dream included her belief that men have a sensitive and responsive penis that affords a clear pleasure. She did also note that this was her phantasy[3] and she did not know whether it was necessarily like this for men. She continued that she did believe that

'their sexual response is not wishy-washy like women's … sometimes there, sometimes not'. (p. 718) In Sarah's psychic landscape, she attributed the penis with the power to desire and without it she was left passively waiting to be desired. We also see that she was burdened with holding rigidly normative views of the meaning of masculinity and femininity and maleness and femaleness.

For Sarah, her dream work revealed her wish to have everything, including the male sexual organ which stood in her psychic economy for the possibility of completeness and certitude. At the same time, on another level, in her dreams, the male organ never was quite right, never quite worked as it should, exposing her anxiety that her appropriation was feeble. Sarah appeared to equate in her imagination maleness with potency, adequacy, and agency. She wanted to embody these qualities, but they were elusive to her. Without these qualities, she felt that she was left with nothing to counter her inescapable belief that she was odd, weird, and damaged both in body and mind. Sarah used her anxieties about her body to communicate her central conundrum, that is, her anxiety in relation to her primary internal object. She believed the other/the analyst did not freely recognize her as a person of worth and value; a person worthy of love. She believed that she could only receive attention and a conditional form of love if she used coercive means to bind the other to her. In her phantasy world she believed that if she were perfect the other would recognize and love her. Sarah's unconscious solution to this dilemma was to use both her body (for example, her vaginal reactions) and genital imagery in dreams as an attempt to position herself as omnipotently 'complete'. The eroticization of the phallus represented a defence against her sense of inadequacy and damage as well as an attempt to mark her difference/to deidentify from her maternal object. She appeared to employ a binary set of gendered assumptions whereby male genitalia (and, by association, men) are given privileged status in her (unconscious) mind; thereby here we see how a normatively gendered set of power relationships is reflected and perpetuated in the unconscious mind.

Discussion: Phantasies, Metaphors, and Conceptual Language

Sarah's anxieties regarding identity and self-esteem are played out on a backdrop of gendered evaluative commitments with the body recruited into giving voice to this internal scenario. Her phantasies reflect an unequal given distinction between the sexes that safeguards a patriarchal approach to symbolization. Yet, the body also functions/is constructed not merely as a passive receptacle but as a corporeal force that carries the possibility of arousal and excitement. In Sarah's phantasy world, her body was employed to sign her incompleteness (her sense of lack) and her belief that, if she had everything, she would be complete (and desirable and able to desire). Sarah's confusion and anxiety about her body, her graphic bodily imagery employed to communicate her anxieties, and her idea of a naturalized body used to code socially normative gender and sexual (p. 719) representations, invite interpretative accounts of penis envy and symbolic femininity as

lack. If we stopped at this point, however, we would be as blind as Oedipus, enacting a transference/countertransference to confirm her anxieties that there was no way out of her encapsulated world. I suggest everyone, like Sarah, uses consciously and unconsciously socially normative gender and sexual representations to elaborate our subjective experience.

Laplanche (1997) offers one of the more compelling accounts of how the other is incorporated into the development of mind. His general theory of seduction reverses the Freudian sources of desire by arguing that the child's unconscious and sexuality are formed by the unconscious desire of the other (primary relationship). While caring and attending to her infant's needs (feeding, soothing, bathing, and so on), the mother/primary carer transmits unconsciously enigmatic messages which engender desire, bodily phantasies, and identifications in the infant. This process inevitably generates an 'excess' of meaning; meaning that goes unrepresented and is repressed.[4] This primal seduction by the (m)other that forms the unconscious is not primarily a fantasy (as in the Freudian Oedipus complex) but represents a real experiential situation although not an event-based encounter (as in Freud's seduction theory). It is not only Sarah's but everyone's experience that unconscious messages from the other/culture shape our psychic life and our way of being in the world.

For me, a problem arises when, at times, psychoanalytic theorizing can sit too experientially close to our enigmatically influenced gendered unconscious minds. As stated, Sarah appeared to hold no positive conception of femininity; any sense of agency in her mind was located in the symbolic masculine. Similarly, Marilla Aisenstein (2012), for example, when describing an identification that instils an agency of learning in women, relies on negating but also endorsing language normatively associated with men and the masculine; she termed this identification as virile but non-phallic (see Gyler 2017 for further discussion of this problematic). So, I suggest that this is not only a problem with which Sarah struggles but also it captures a problem for psychoanalysts and psychoanalytic theorizing. I want to emphasize that these issues affect, often in out-of-awareness ways, how we think, especially how we listen, think, and speak in the consulting room; that is, our personal meanings and preferred theories potentially impact the transference/countertransference encounter (Hansell 2011; Kulish 2010). Psychoanalysts are required to hold their explicit and implicit theories and their subjective clinical encounters in a creative tension that values both the feminine and masculine positions in order to construct a third position from which to interpret. Glocer Fiorini (2010) argues that metatheories of the universalization of the feminine as other, the place of lack in the riddle of sexual difference, carry risks. Depending on the implicit theory being employed by a psychoanalyst, the gendered unconscious of the analysand can be understood at the wrong level or misunderstood, for example in relation to women's jealousy and rivalry (p. 720) with men. Emilce Dio Bleichmar asks: 'Is it literally envy of the penis as a male genital organ, or is it symbolic materiality of the social difference between the genders?' (2010: 197). Putting this problematic another way: on the one hand, hyperfeminized states (hysterical and histrionic states) can go

unquestioned in women, and on the other, women's ambitions can be interpreted solely in relation to penis envy rather than a wish to expand possibilities for themselves.

Gender Variance

Trans and queer gender theory represents a site of complex intersections between psychoanalysis and cultural theory and moves away from challenging binary constructions in terms of the power relations between female/male, feminine/masculine oppositions. It deconstructs binary thinking from a different trajectory—it challenges the naturalized relationship between the anatomical body and gender identity and in doing so questions any universalizing conceptions of woman/man. It produces a discourse that Goldner described as gender in 'free fall' (2011: 159). The trans cultural gender literature is diverse, multilayered, and bursting with contested interpretations and it is beyond the scope of this chapter to review this literature. I shall, however, call attention briefly to one debate that has resonance for clinical practice. The wrong-body theory[5] versus the beyond-the-binary model (that is, no recognition given to one or other categories) debate goes to the heart of theorizing identity issues—for example is gender identity innate? Are sex and gender socially constructed?[6] Talia Bettcher argues that the beyond-the-binary model and the wrong-body theory are both embedded in the conventionally understood gender categories:

> In the beyond-the-binary model, to say that trans people are marginal with respect to the binary is to locate them in terms of the categories 'man' and 'woman' as dominantly understood. If trans bodies can have different resistant meanings, the decision to say of those bodies that they are 'mixed' or 'in between' is precisely to assume a dominant interpretation. So, the problem is not the rigidity of the binary categories but rather the starting assumption that there is only one interpretation in the first place (the dominant one). Similarly, in the wrong-body model, to become (p. 721) a woman or a man requires genital reconstruction surgery as a correction of wrongness. But this is to accept a dominant understanding of what a man or a woman is. (2014: 390)

For Bettcher, both accounts are underpinned by normative assumptions that leave the logic of the gender binary intact.

In contrast to the cultural theory literature, the psychoanalytical contribution is relatively limited on the question of transgender. Following Stoller (1968) and our history of theoretical and clinical errors regarding same-sex desire, this literature does not usually equate gender variance with psychopathology. Expressions of gender and sexuality are no longer viewed as giving a privileged view of psychic functioning, but in some circumstances may be recruited to communicate mental distress and disorder (Saketopoulou 2011). To give an example from the recent psychoanalytic literature: Avgi Saketopoulou (2014) describes the clinical case of Jenny, a five-year-old trans girl. Unlike Sarah, described above, who experienced her form of subjectivity as wrong, Jenny experienced her gender/body misalignment as wrong and unmirrored and misrecognized

by those in her environment. Jenny presented herself as female from a very early age (two years old) and became dysregulated and angry if referred to by her birth name, Johnny. According to Saketopoulou, by age five Jenny was experiencing a seriously depressed mood necessitating psychiatric intervention and bodily dysphoria with a near delusional belief about her anatomical body. For Jenny, her gender identity and natal body did not align according to her perception of family and societal expectations and norms, producing a sense of excruciating shame, sadness, and guilt. Saketopoulou reported that Jenny said: 'Dad thinks I'm a boy … sometimes I wear boy clothes, so his heart doesn't keep breaking' (2014: 774). The reader is left to ponder the nature of the enigmatic messages unconsciously transmitted to Jenny in her early life that now animate her psychic reality—but I agree that Saketopoulou is right to 'deliberately eschew the much-debated quest for etiological factors' (2014: 775). While psychoanalytical methodology may assist in (co-)creating a personal reconstructive narrative, this is not in the same category as causal discourses. In this case study, the realness of Jenny's anguish exposes the psychic burden when inner reality and the normative gendered binary system collide. For Jenny, a naturalized alignment should exist between gender identity and the natal body.

Conceptually, this case complicates how we theorize the body (contentious because the body has always been a fraught place for feminism). As referred to earlier, it is now usual to see the body as both made and given; both mediated and immediate; both a corporeal force and an idea; both never prediscursive and a bodily force that exceeds symbolization (Dimen 2014). Saketopoulou appears to take a different step which is not uncommon in the psychoanalytic literature theorizing trans experience, in conceptualizing the body as naturalized body (see Dimen 2014). First, there appears an acceptance of a naturalized link between the body and gender, that is, one's subjective sense of gender identity should align with one's anatomical/natal body. Second, this link is denaturalized so that, as Saketopoulou writes, there is an 'undoing of the notion (p. 722) that her (Jenny's) body delivered gender's verdict' (2014: 788). The body as naturalized, denaturalized, and often renaturalized, nevertheless remains conceived in relation to the binary modality of thought. As Bettcher put it, greater freedom of gender representation and expression requires the 'starting assumptions' of the gender binary to be reframed to allow for the meaning of the categories of women and men to be decentred, along with the relation between gender identity and the anatomical body.

In future accounts of decentred and deconstructed subjectivity, the task continues to be the development of frameworks that permit a representational space for the body, and the meaning of procreative body in psychic life, mediated through phantasy, desire, and affect without essentializing and universalizing it. Rosemary Balsam (2012) has called attention to the elision of the pregnant body in psychoanalytic discourse which has resulted in diminished accounts of the representation of women. While psychoanalytic theory and practice do not map together in any one-to-one correspondence, I suggest that this representational gap will influence how an analysand's associative speech is registered, organized, and interpreted by the analyst. I have argued elsewhere (Gyler 2010) that if Freud had recognized the vagina and its meaning, his Oedipal theory and especially his

notion of the female Oedipal complex would have been overturned. Representations of female genitalia and of procreativity challenge formulations of the feminine position as only conceived as castrated/lack/absence and lead to the need to include a different version of femininity with bodily based meanings (see Birksted-Breen 1999).

Concluding Thoughts

Cultural gender theories opened the way for critical appreciation of the gendered discourses of psychoanalysis as well as expanded our thinking about the possible repertoire of individual gender variations. My starting point for this chapter asked: does the concept of gender continue as a critical category for contemporary clinical psychoanalysis? I suggested that in exploring this question there are two contemporary developments that might lead psychoanalysts to consider that gender as a critical category is no longer as cogent as it once was. The first hinges on the belief that the deconstructive and reconstructive endeavours of past decades have sufficiently troubled and undone the heterosexist and patriarchal assumptions and values embedded in earlier theories. The other direction relates to the increasing attention given to working with analysands where archaic states of mind predominate their ways of being in the world.

On the other hand, I suggest that gender enters clinical practice through the individual gendered subjectivity of the analyst and the analysand and through gendered discourses. The analysts bring conscious and unconscious assumptions, biases, and theoretical commitments that influence their receptivity and countertransference responses and inform their style of participation. For analysands, questions of gender identity may or may not be the site of maximal anxiety, but gender anxieties may be expressed, for (p. 723) example, through the inevitable anxiety and loss associated with gendered subjectivity, the social suffering of individuals with variant and resistant gender identities, and the recruitment of gender expressions and gendered phantasies to communicate unmourned mental states and distress. Nevertheless, the question of gender is no longer a self-evident arbiter of psychopathology for contemporary practice.

As we grapple with these questions we, as individuals, can think in terms of gender categories or resist notions of categorization, especially normative versions. Problems can arise when ideas about gender carry rigid and fixed meanings which dispense negative and destructive burdens. I suggest that (leaving aside the debate as to whether he was describing or/and prescribing) Freud's ambivalently held conventional equation of male equals strong and female equals weak continues to reverberate in the psyches of men and women, either as conformity or resistance or in mostly combined and contradictory ways. Perhaps it is inevitable that, regardless of our conscious intentions, the orthodoxy of this equation unavoidably echoes in the consulting room both in the conscious and unconscious minds of patients and analysts alike because we all frequently employ socially normative gender and sexual representations to organize and elaborate our subjective experience. This constrains the ways we see, listen, think, and talk (that is, we construct meaning through frameworks; there is no view from 'nowhere'). Everyone is

challenged to remain in an open-ended and uncharted space which, at times, feels unbearably unmoored. Chodorow puts it this way: 'In our [psychoanalysts'] thinking about gender, as in our thinking about other aspects of psychological functioning, it is always so much easier to have set theories to bring to bear on what we hear, rather than stay with uncertainty' (1995: 298). Our knowledge of individuals is not exemplary. While the challenge is to remain open to not knowing, we also need to remember that the gendered binary of Western thought inevitably exerts its pull and problematizes the representational space for the feminine, same-sex desire, and the body.

Psychoanalysis offers rich descriptive and experience-near accounts of the workings of the mind and meaning-making. The Western gendered binary structure of thought shapes the phantasies, metaphors, and ideas used to make meaning. In theorizing, it can be a complicated task to disentangle the different levels of understandings; phenomenological accounts and conceptual abstraction can be conflated. Analysands' unconscious phantasies which might employ normative social values and family arrangements can be transcribed literally into conceptual and theoretical language. I suggest that the effect of these discourses is to both produce and be productive of heteronormative frameworks. To probe these discourses is not to elide the question of difference but to acknowledge, as Scott (1999) argues, the need for systematic criticism of the operations of categorical difference with its hierarchical inclusions and exclusions and to refuse any 'truth' implied in these constructions not in the name of equality based on sameness but rather an equality that recognizes differences.

In this chapter, I started with the question as to whether gender had been sufficiently deconstructed in our discourses to allow the contemporary clinician to be less 'troubled' by questions of gender. I have attempted to indicate that the question of gender needs to be contextualized in the broader preoccupations for psychoanalysis. I suggest the (p. 724) question of gender remains of crucial concern for representing human possibilities and all three of Kristeva's attitudes continue to require our ongoing attention. To borrow from Donna Haraway a phrase used in a different context, we need to 'stay with the trouble' (2016) to recognize with pleasure rather than fear the otherness of difference.

References

Aisenstein, M. (2012). 'A particular form of paternal identification in women'. *Psychoanalytic Quarterly 81*(1): 27–37.

Aisenstein, M. (2015). 'The question of the Father in 2015'. *Psychoanalytic Quarterly 84*(2): 351–362.

Appignanesi, L. and Forrester, J. (1993). *Freud's Women*. London: Virago Press.

Balsam, R. H. (2012). *Women's Bodies in Psychoanalysis.* New York: Routledge.

Beauvoir, S. de (1949/1972). *The Second Sex*, trans. H. M. Parshley. Harmondsworth: Penguin.

Benjamin, J. (1996). 'In defense of gender ambiguity'. *Gender and Psychoanalysis: An Interdisciplinary Journal 1*(1): 27–44.

Bettcher, T. (2014). 'Trapped in the wrong theory: Rethinking trans oppression and resistance'. *Signs: Journal of Women in Culture and Society 39*(2): 383–404.

Bion, W. R. (1959/1967). 'Attacks on linking'. In *Second Thoughts: Selected Papers on Psychoanalysis* (pp. 93–109). London: Maresfield Library.

Bion, W. R. (1962a). *Learning from Experience*. London: Heinemann Medical.

Bion, W. R. (1962b/1967). 'A theory of thinking'. In *Second Thoughts: Selected Papers on Psychoanalysis* (pp. 110–119). London: Maresfield Library.

Bion, W. R. (1970). *Attention and Interpretation.* London: Maresfield Library.

Bion, W. R. (1992). *Cogitations.* London: Karnac Books.

Birksted-Breen, D. (1999). 'Holding the tension'. In D. Bassin (ed.), *Female Sexuality: Contemporary Engagements*. New Jersey: Jason Aronson Inc.

Botella, C. and Botella, S. (2005). *The Work of Psychic Figurability: Mental States without Representation.* London: Routledge.

Butler, J. (1990). *Gender Trouble: Feminism and the Subversion of Identity*. New York: Routledge.

Butler, J. (2004). *Undoing Gender*. New York: Routledge.

Chodorow, N. (1978). *The Reproduction of Mothering: Psychoanalysis and the Sociology of Gender*. Berkeley, CA: University of California Press.

Chodorow, N. (1994). *Femininities, Masculinities, Sexualities: Freud and Beyond*. London: Free Association Books.

Chodorow, N. (1995). 'Multiplicities and uncertainties of gender: Commentary on Ruth Stein's "Analysis of a Case of Transsexualism" '. *Psychoanalytic Dialogues 5*(2): 291–299.

Chodorow, N. (2012). *Individualizing Gender and Sexuality: Theory and Practice.* New York: Routledge.

Corbett, K. (2008). 'Gender now'. *Psychoanalytic Dialogues 18*(6): 838–856.

Dimen, M. (2014). 'Both given and made: Commentary on Saketopoulou'. *Journal of the American Psychoanalytic Association 62*(5): 807–813.

Dio Bleichmar, E. (2010). 'The psychoanalyst's implicit theories of gender'. In L. Glocer Fiorini and G. Abelin-Sas Rose (eds), *On Freud's 'Femininity'.* London: Karnac.

Elise, D. (2008). 'Sex and shame: The inhibition of female desire'. *Journal of the American Psychoanalytic Association 56*(1): 73–98. (p. 725)

Eshel, O. (2017). 'From extension to revolutionary change in clinical psychoanalysis: The radical influence of Bion and Winnicott'. *Psychoanalytic Quarterly 86*(4): 753–794.

Flax, J. (2006). 'Masculinity and its discontents: Commentary on Reichbart and Diamond'. *Journal of the American Psychoanalytic Association 54*(4): 1131–1138.

Freud, S. (1905). *Three Essays on the Theory of Sexuality.* In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 7 (pp. 123–243). London: Hogarth Press.

Freud, S. (1925). 'Some psychical consequences of the anatomical distinction between the sexes'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 19 (pp. 241–258). London: Hogarth Press.

Freud, S. (1926). 'The question of lay analysis'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 20 (pp. 177–250). London: Hogarth Press.

Freud, S. (1931). 'Female sexuality'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 21 (pp. 221–243). London: Hogarth Press.

Freud, S. (1933). *New Introductory Lectures on Psychoanalysis*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 22 (pp. 1–182). London: Hogarth Press.

Gagnier, R. (2017). 'A symbiological approach to sex, gender and desire in the Anthropocene'. *Angelaki: Journal of the Theoretical Humanities 22*(1): 11–21.

Glocer Fiorini, L. (2010). 'The analyst's meta-theories concerning sexual difference and the feminine'. In L. Glocer Fiorini and G. Abelin-Sas Rose (eds), *On Freud's 'Femininity'*. London: Karnac.

Glocer Fiorini, L. (2017). *Sexual Difference in Debate.* London: Karnac.

Goldner, V. (2011). 'Trans: Gender in free fall'. *Psychoanalytic Dialogues 21*(2): 159–171.

Green, A. (1995). 'Has sexuality anything to do with psychoanalysis?' *The International Journal of Psychoanalysis 76*(5): 871–883.

Gyler, L. (2010). *The Gendered Unconscious: Can Gender Discourses Subvert Psychoanalysis?* London: Routledge.

Gyler, L. (2017). 'From symptom to symbolization of receptivity: A Girl's psychoanalytic journey'. *Angelaki: Journal of the Theoretical Humanities 22*(1): 35–47.

Hansell, J. (2011). 'Where sex was. There shall gender be? The dialectics of psychoanalytic gender theory'. *Psychoanalytic Quarterly 80*(1): 55–71.

Haraway, D. (2016). *Staying with the Trouble: Making Kin in Chthulucene*. Durham, NC: Duke University Press.

Harris, A. (1997). 'Aggression, envy and ambition: Circulating tensions in women's psychic life'. *Gender and Psychoanalysis 2*(3): 291–325.

Harris, A. (2005). *Gender as Soft Assembly*. Hillsdale, NJ: The Analytic Press.

Horney, K. (1925/1999). 'The flight from womanhood: The masculinity complex in women, as viewed by men and women'. In R. Grigg, D. Hecq, and C. Smith (eds), *Female Sexuality: The Early Psychoanalytical Controversies* (pp. 107–121). London: Rebus.

Isaacs, S. (1948). 'The nature and function of phantasy'. *The International Journal of Psychoanalysis 29*: 73–97.

Jones, E. (1927). 'The early development of female sexuality'. In R. Grigg, D. Hecq, and C. Smith (eds), *Female Sexuality: The Early Psychoanalytical Controversies*. London: Rebus Press.

Joseph, B. (2013). 'Here and now: my perspective'. *The International Journal of Psychoanalysis 94*(1): 1–5. (p. 726)

Klein, M. (1928). 'Early stages of the Oedipus conflict'. In M. Klein, *Love, Guilt and Reparation and Other Works 1921–1945*. London: Karnac.

Korbivcher, C. (2014). *Autistic Transformations: Bion's Theory and Autistic Phenomena.* London: Karnac.

Kristeva, J. (1979/1986). 'Women's time'. In T. Moi (ed.), *The Kristeva Reader* (pp. 187–213). Oxford: Basil Blackwell.

Kulish, N. (2010). 'Clinical implications of contemporary gender theory'. *Journal of the American Psychoanalytic Association 58*(2): 231–258.

Laplanche, J. (1997). 'The theory of seduction and the problem of the other'. *The International Journal of Psychoanalysis 78*(4): 653–666.

Mitchell, J. (1974). *Psychoanalysis and Feminism*. Middlesex: Allen Lane.

Mitchell, J. (2015). 'Debating sexual difference, politics, and the unconscious with discussant section by Jacqueline Rose'. In R. Duschinsky and S. Walker (eds), *Juliet Mitchell and the Lateral Axis: Twenty-First-Century Psychoanalysis and Feminism.* New York: Palgrave Macmillan.

Mitchell, S. A. (1996). 'Gender and sexual orientation in the age of postmodernism: The plight of the perplexed clinician'. *Gender and Psychoanalysis: An Interdisciplinary Journal 1*(1): 45–74.

Money, J. (1973). 'Gender role, gender identity, care gender identity: Usage and definition of terms'. *Journal of the American Academy of Psychoanalysis 1*(4): 397–402.

Oakley, A. (1972). *Sex, Gender and Society*. London: Temple Smith.

O'Shaughnessy, E. (2013). 'Where is here? Where is now?' *The International Journal of Psychoanalysis 94*(1): 7–16.

Perelberg. R. J. (2013). 'Paternal function and thirdness in psychoanalysis and legend: Has the future been foretold'. *Psychoanalytic Quarterly 82*(3): 557–585.

Person, E. S. (1999). *The Sexual Century*. New Haven, CT: Yale University Press.

Roussillon, R. (2011). *Primitive Agony and Symbolization.* London: Karnac.

Rubin, G. (1975). 'The traffic in women: Notes of the "political economy" of sex'. In R. Reiter (ed.), *Toward an Anthropology of Women.* New York: Monthly Review Press.

Saketopoulou, A. (2011). 'Minding the gap: Intersections between gender, race, and class in work with gender variant children'. *Psychoanalytic Dialogues 21*(2): 192–209.

Saketopoulou, A. (2014). 'Mourning the body as bedrock: Developmental considerations in treating transsexual patients analytically'. *Journal of the American Psychoanalytic Association 62*(5): 773–806.

Scott, J. W. (1999). *Gender and the Politics of History*. New York: Columbia University Press.

Scott, J. W. (2011). *The Fantasy of Feminist History*. Durham, NC/London: Duke University Press.

Showalter, E. (1985). *The Female Malady: Women, Madness and English Culture, 1830–1980*. London: Virago Press.

Stoller, R. (1968). *Sex and Gender.* New York: Jason Aronson.

Tustin, F. (1988). 'Psychotherapy with children who cannot play'. *The International Review of Psycho-Analysis 15*(1): 93–106.

Unger, R. K. (1979). 'Toward a redefinition of sex and gender'. *American Psychologist 34*: 1085–1094.

Young-Bruehl, E. (1996). 'Gender and psychoanalysis: An introductory essay'. *Gender and Psychoanalysis 1*(1): 7–18.

Notes:

(^{1}) [Eds: 'Enigmatic signifier' is a term (taken by Jean Laplanche from Jacques Lacan) to mean language which transmits meaning unconsciously, such meaning being largely composed of sexual desire that has been repressed.]

(^{2}) [Eds: For an extended discussion of Freud's conception of sexuality, see Sandford (this volume).]

(^{3}) The 'ph' spelling of phantasy is used to denote the Kleinian concept which refers to the 'primary content of unconscious mental processes' (Isaacs 1948: 81). This definition extends the notion of fantasy as wish-fulfilment.

(^{4}) [Eds: For a critical expository account of Laplanche's theory of enigmatic signifiers, see Fletcher, J. (2007). 'Gender, sexuality and the theory of seduction'. In L. Braddock and M. Lacewing, *The Academic Face of Psychoanalysis* (pp. 224–40). London: Routledge.]

(^{5}) According to Bettcher (2014), the wrong-body theory can be schematically conceived by a weaker version proposing proper alignment with an innate gender identity and a stronger version suggesting 'real' sex determined by gender identity. Nevertheless, the implications drawn from this theory for a lived life are mixed: for example, some cultural theorists argue that body reassignment surgery is medicalizing the body, with all the associated dangers of historical interventions into the body/sexuality to 'naturalize' and constrain difference and otherness; while others see the 'wrong-body theory' as a possible route to recognition that bypasses experience of doubt and blame and accepts and supports medical and surgical interventions.

(^{6}) Bettcher (2014) makes the critical point that social constructions are no less 'real' than apparent 'givens'.

Louise Gyler

The Australian Psychoanalytical Society, Australia

Oxford Handbooks Online

Political Philosophy in Freud: War, Destruction, and the Critical Faculty 🔓

Abstract and Keywords

Freud's reflections on death and the death drive form a proper part of political philosophy. If the idea of a self-governing polis requires that citizens maintain a set of social bonds, then the threat to social bonds also threatens the possibility of self-rule. Freud identifies a destructive impulse or drive as part of the human psyche, clarifying that it has the specific capacity to destroy social bonds. This chapter considers the importance of ambivalence as a permanent feature of love relations and those social relations that form the basis of political life. It argues that a political effort to counter destruction must call upon the resources of melancholia, including the manic refusal of tyranny and a normalizing reality principle. Defending the idea of militant pacifism, the argument here suggests that aggression can be—and must be—part of the strategy against violence and the sustainable future of political life.

Keywords: death, death drive, love, Eros, politics, political philosophy, melancholia, mania, pacifism, war

> I fear I may be abusing your interest, which is after all concerned with the prevention of war and not with our theories. Nevertheless I should like to linger for a moment over our destructive drive, whose popularity is by no means equal to its importance.
>
> Freud to Einstein, 1932

Introduction

In general, Freud is not considered to be a political philosopher although some of his insights have been brought into philosophical positions that seek to address political issues. Sometimes political metaphors turn out to have special importance to certain views he has enunciated. So, for instance, his work on repression makes use of the language of censorship, though, strictly speaking, there is no state or institution that directs censorship when ideas or wishes are censored. The ego is occasionally described as an 'institution', giving it a social status and potential political significance. On several occasions, he remarks on the dangers of nationalism, the phantasmatic character of citizenship, the problem of authority, the origins of violence and war, and the prospects for peace. At work here is an operative metaphor of the psyche as a legal system unto itself, passing judgement, asserting authority, banishing foreign elements, and dealing out forms of punishment. In his 'Thoughts for the times on war and death', written in 1915 and in the midst of the First World War, Freud reflected on the bonds that hold a community together as well as the destructive powers that break those bonds (Freud 1915b). By the time he develops the death drive, first in 1920 and then more fully in the following decade, he becomes increasingly concerned with the destructive capacities of human beings. What he calls sadism, aggression, and destructiveness come to be primary representatives of the death drive, which receives its most mature formulation in *Civilization and its Discontents* in 1930. The 'unconquerable' drive takes on a new form as Freud develops a dualistic metaphysics, counter-posing Eros, the force that creates ever more complex human bonds, to Thanatos, the force that breaks them down (1930: 86). A durable political form presumes that social bonds can remain relatively in place, but how, then, do polities deal with the destructive force that Freud describes?

Political philosophy traditionally seeks to understand concepts such as justice, freedom, and equality, but usually only in the context of a broader question: what is the best way for humans to govern themselves in common? It presumes that polities are established and maintained, and that laws and norms exist to regulate these forms of shared existence or common life. So when we ask about that best way for human beings to govern themselves, and we look for state structures or forms of governance that will answer that question, we make several key assumptions. First, we presume that humans are in some sense disposed to govern themselves in common, or that they can become so disposed. Either they have a concern for others, are part of a general will, have a concern with what is best for all, or have a social and political sense of having to find a way of being one single self among many, or they can be educated or compelled to develop an orientation of this kind. We tend to presume as well that the particular relations established within any given polity are oriented towards the continuation of the polity except in those cases where the polity proves to be so thoroughly unworkable or unjust that it has to be dissolved in favour of another regime. If Aristotle is right that human beings are political animals, that they only properly know and express themselves as human in the context of political life, then it follows that the particular form that human

Subscriber: Tulane University; date: 15 January 2019

relations assume within a polity directly bears on the question of whether what is human about them can know and express itself within that polity. For humans to govern themselves and to do that together seems to presuppose that they are already related to one another enough for that prospect to be realized, or that they can become related to one another in the course of governing. The possibility of realizing political life not only assumes a shared life, but a set of social bonds that reflects a commitment to perpetuating that shared life, finding a mode of governance, and participating in the process of governing themselves.

The effort to understand how a social bond is forged and maintained leads us to Freud's view that it is Eros that compels us to create and sustain human bonds, even to elaborate increasingly complex human bonds. Thus, according to that view, polities that depend on the possibility of continuing and expanding increasingly complex human ties may be said to rely upon Eros in some way. If they presume that people will be oriented towards maintaining and cultivating such ties, then those polities—and the political philosophies by which they are conceptually grasped—presume the workability of those bonds and must, as a result, anticipate their breakability as well. In his reflections on war, for instance, the breakability of those bonds becomes for Freud a pressing matter for reflection.

Freud on War

Freud's reflections on the First World War led to successive insights on destructiveness. In 1915, Freud has not yet introduced the notion of the death drive—which will have as one of its primary aims the deterioration of social bonds—but he does register an impression of overwhelming and unprecedented human destructiveness in his time:

> … the war in which we had refused to believe broke out and brought—disillusionment. Not only is it more bloody and more destructive than any war of other days, because of the enormously increased perfection of weapons of attack and defence; it is at least as cruel, as embittered, as implacable as any that has preceded it. It disregards all the restrictions known as International Law, which in peace-time the states had bound themselves to observe; it ignores the prerogatives of the wounded and the medical service, the distinction between civil and military sections of the population, the claims of private property. It tramples in blind fury on all that comes in its way, as though there were to be no future and no peace among men after it is over. It cuts all the common bonds between the contending peoples, and threatens to leave a legacy of embitterment that will make any renewal of those bonds impossible for a long time to come

> (Freud 1915b: 278–9).

Freud's remarks are noteworthy for many reasons, chief among them the sense of shift in the history of destructiveness: destructiveness has not been known quite like this before. Although the development of new weapons made the destruction greater than in other wars, the level of cruelty strikes Freud as the same, suggesting that the problem is not that humans have become more cruel, but that technology has allowed that cruelty to produce greater destruction than before. A war without those weapons would cause less destruction but would engage no lesser amounts of cruelty. So if we are tempted to say that cruelty is itself augmented by technology, Freud appears to resist that view: destruction takes on new and historically variable forms, but cruelty remains the same. Thus, human cruelty alone does not account for all destructiveness—technology exercises its agency as well. But the distinctly human capacity for destructiveness in human beings follows from the ambivalent psychic constitution of the human subject. The question of what can be done to check destructiveness thus engages ambivalence and technology, especially in the context of war.

Although it is commonly supposed that war-making is the specific activity of nations, the blind rage that motivates war destroys the very social bonds that make nations possible. Of course, it can fortify the nationalism of a nation, producing a provisional coherence bolstered by war and enmity, but it also erodes the very bonds that make politics possible. Not every polity is a nation state, and the nation cannot be reduced to the polity. What Freud's analysis allows us to understand is that if a polity depends on accords with other polities, or if it depends on laws, including those laws that govern military action, then war imperils those accords and those laws. In fact, war tends to be oblivious to key dimensions of governance such as international law and human rights, the rights of nations, the obligation to protect the lives of civilians against harm, and rights to private property—to place them in suspension, or to calculate on what forms of criminality will be tolerated by the international community. The power of destruction unleashed by war breaks social bonds and produces anger, revenge, and distrust ('embitterment') such that it becomes unclear whether reparation is possible, undermining not only those relations that may have been built in the past, but also the future possibility of peaceful coexistence. Although Freud is clearly reflecting on the First World War in his remarks here, he is also making claims about war in general: war 'tramples … on all that comes its way'. Here he is suggesting that breaking down the barriers that keep inhibitions in place is, in fact, one aim of war—military personnel have to be given licence to kill. Whatever the explicit strategic or political aims of a war may be, they prove to be weak in comparison with the aims of destruction; what is destroyed first in war are the very restrictions imposed on destructive licence. If we can rightly speak about the unstated 'aim' of war, it is neither primarily to alter the political landscape nor to establish a new political order but, rather, to destroy the social basis of politics itself. Of course, such a claim may seem overstated if we believe, for instance, in just wars, wars waged against fascist or genocidal regimes in the name of democracy. But even then, the explicitly stated goal of war-waging and the destructiveness unleashed by war are never quite the

same. Even the so-called 'just war' runs the risk of a destructiveness that exceeds its stated aims, its deliberate purpose.

This last claim might at first appear perplexing since governments do give reasons for going to war: the expansion or recovery of territory and natural resources, the defences of people or land, the restoration of honour, patriotism, and the like. And yet, Freud makes two claims that call this account into question. First, whatever the public and stated aims of war may be, another aim is always also at work, one that Freud refers to here as 'a blind fury'. Second, this fury, motivating and even unifying a people or a nation at war, also tears the people and the nation apart, working against whatever intentional, self-preserving, or self-enhancing aims they may have. The aims of this sort of rage include first and foremost overcoming existing inhibitions and restrictions imposed on destruction itself, breaking social bonds—understood in part as blocks against destruction—in favour of increased destructiveness, reproducing destruction into the foreseeable future, which may turn out to be either a future of destruction or a way of destroying the future itself. It is from within the stated, local, and provisional aims for war-making that another aim or indeed 'a drive' can take hold, a destructiveness without limit. Even as a group or a nation may achieve temporary cohesion in war, rallying behind its explicit aims to defend the country or to destroy the enemy, something can form—or take hold—within that rallying that exceeds any of those explicitly acknowledged aims, not only breaking the social bonds of the groups waging war but those targeted by war as well. The idea of 'blind fury' that Freud takes from Greek tragedy prefigures what he will come to call the death drive just five years later. Already in 1915, what concerns him is the power that the death drive assumes once it is amplified with destructive technology to wreak destruction across the world, and to destroy the very social bonds that have the power to keep destructiveness in check. By 1930, Freud will become more explicitly concerned with the possibility of genocide, as evidenced in *Civilization and its Discontents*. There he writes:

> The fateful question for the human species seems to me to be whether and to what extent their cultural development will succeed in mastering the disturbance of their communal life by the human instinct (*Trieb*) of aggression and self-destruction [*Agressions und Selbstvernichtungstrieb*]. It may be that in this respect precisely the present time deserves a special interest. Men have gained control over the forces of nature to such an extent that with their help they would have no difficulty in exterminating one another to the last man
>
> (Freud 1930: 145).

He appends a line to that paragraph in 1931 that calls upon 'eternal Eros … to assert himself in the struggle with his equally immortal adversary', noting that no one can foresee how successful that effort will be (Freud 1930: 145). Freud is clearly looking to find a possibility in the life of the drives to counter the horrific destructiveness that he saw in the First World War and that he senses is returning to Europe in greater proportions in the 1930s. Freud does not turn to history or to empirical examples in his

effort to understand destructiveness, but to what he calls the 'drives'—a move that seems speculative at best. So why look to the life of the drive? For Freud, the conscious reasons for acting that groups give to themselves are not the same as the underlying motivations that guide their action. As a result, reflection on how best to avert destruction must do something other than provide an argument acceptable to rational thought—it must somehow appeal to the drive, or find a way of working with—and against—that propulsive destructiveness that can lead to war.[1]

One sceptical position towards drive theory results from a mistaken translation of 'instinct' for '*Trieb*' in Freud's writing. Although *Instinkt* and *Trieb* are both used in Freud's work, the latter appears more often, and the death drive (*Todestrieb*) is never 'the death instinct'. The Strachey translation of the *Complete Works* has consistently translated both terms as 'instinct', giving rise to a biologistic understanding of the term in English-language literature and, in some cases, a view that drives, in Freud, follow a form of biological determinism. Freud makes clear in an essay entitled 'Instincts and their Vicissitudes' (*Triebe und Triebschicksale*) but that should be called 'Drives and their Destinies' that the drive (*Trieb*, meaning 'push') belongs neither exclusively to the realm of biology nor to a fully autonomous psychic domain; rather, it functions as a threshold concept (*ein Grenzbegriff*) between the somatic and ideational spheres (1915a).[2] Over and against a biologistic reduction of the drive, he argues at one point that 'drives' belong to the poetic dimension of his own work.[3] In fact, in *Beyond the Pleasure Principle*, he affirms that figurative language is peculiar to psychology and to scientific language itself, suggesting that efforts to provide biological explanations might not be altogether successful in purging figurative language from their operation. Freud's point, however, is to call into question the idea that drives are biological sources of action that act in unknown and deterministic ways. It seems much more plausible to conclude that drives function as powerful motivations that become manifest in feeling and actions that cannot be fully understood through conscious reflection alone.

The Death Drive

Until 1920, Freud maintained that psychic life was governed by pleasure, sexuality, or libido, and it was only when he encountered forms of war neurosis that he began to consider that there were symptoms characterized by compulsive repetition that could not be explained by wish-fulfilment or a drive towards gratification. So it was in the wake of war that Freud began to formulate the death drive, prompted as well by his consideration of forms of destructiveness with a repetitive quality that seemed distinct from all erotic satisfaction (what he will later refer to as 'non-erotic aggressivity' in *Civilization and its Discontents* (1930: 120)). In the first formulation of the death drive in *Beyond the Pleasure Principle*, Freud sought to find an explanation for forms of repetitive behaviour that did not appear to be guided by satisfaction. He had encountered patients suffering from war neurosis who relived traumatic scenes of violence and loss in ways that bore no clear resemblance to forms of repetition linked with libido and wish-fulfilment. There appeared to be no satisfaction linked to this kind of uncontrollable repetition of suffering, progressively deteriorating the condition of the patient. At this stage, Freud developed the first version of the death drive according to which the organism seeks a return to its primary inorganic state, a state relieved of all excitation. Indeed, every human organism seeks to return to this origin such that the trajectory of a life turns out to be no more than a 'circuitous route toward death' (1920: 38). [4] As much as there is something in humans that seeks to fulfil wishes and to preserve its own organic life, there is also something that operates to the side of wish-fulfilment, seeking to negate the organic conditions of life, whether that life belongs to another or to oneself.

At this point Freud effectively calls into question one of the key tenets of classical liberal political philosophy, namely, that the primacy of self-preservation functions as a presupposition in the formation of social groups and forms of government. If one were to argue in a Hobbesian vein that self-interest or self-preservation leads humans into war, one would be neglecting the very dimension of destructiveness that Freud seeks to foreground, one that potentially works against both self-interest and self-preservation. Freud's point is important since it would seem that political forms depend upon the preservation of social bonds, so if something works to destroy those bonds, it works to undermine the basis of political community as well. The preservation of social bonds presumes that beings seek to preserve themselves not only as individuals, but as social creatures for whom preservation depends upon ongoing social relations. The workability of this claim presumes, however, that self-preservation is bound up with social relations, that is, that the 'self' who is preserved is defined in fundamental ways by its social relations and not only or primarily as an individual.

If a being whose primary aim is self-preservation has to live with others who have the same primary aim, then how do they seek (a) to secure the preservation of each other's lives, and (b) develop forms of concern for society as a whole? One answer to that dilemma has been to show that society itself secures self-preservation, and so fulfils that primary aim of human life. What difference does it make that Freud now posits another

Subscriber: Tulane University; date: 15 January 2019

tendency within the human psyche that seeks to return it to a time before the individuated life of the human organism? If self-destruction is as equally powerful as self-preservation, what implications does that have for the political forms in which people live? His reflections on destruction focus on the possibility of the destruction of other lives, especially under conditions of war, in which the technology of weaponry amplifies the powers of human destructiveness. Those who suffered war neurosis were living out the psychic consequences of war, but they also became the occasion for Freud to consider how destruction works not only against others, but against oneself. War neurosis continues the suffering of war in the form of traumatic symptoms characterized by repetition. This form of repetition serves no wish-fulfilment, progressively deteriorating the condition of the patient with no end in sight. Freud identifies this as the repetitive character of destruction. In the patient, it eventuates in social isolation; more broadly considered, it not only serves to weaken the social bonds that hold societies together, but also to deteriorate the human organism, taking form as a self-destruction that can culminate in suicide. Libido or sexuality has a reduced or vanishing role in this form of destruction, and the social bonds without which political life proves impossible are shredded in its midst. Of course, there are communities and bonds of solidarity that are formed under conditions of war, apparently vitalized by shared—transitively communicated—aims of destruction. We see this across the political spectrum, social bonds formed in resistance or in opposition, taking up arms to achieve their goals. The more general point, however, is that certain conditions of destruction erode and potentially destroy even those forms of social relations, since trauma disconnects a subject from itself and from the other. When this happens on a large scale, the social basis of political life is imperilled; repair and renewal are indefinitely suspended.

Towards the end of *Beyond the Pleasure Principle*, Freud asserts that every human organism in some sense seeks its own death, and that this tendency cannot be derived from the sexual drives. The evidence for the death drive, he argues, can additionally be found within sexual sadism and, more generally, within the phenomenon of sadomasochism.[5] Although the sexualization of the death drive can subordinate its destructiveness to what Freud regards as the non-destructive aims of sexuality, the death drive can come to predominate—a situation that would be illustrated clearly with sexual violence. Both self-destruction and the destruction of the other are potentially at work within sadomasochism, suggesting for Freud that a drive separate from the sexual drive can nevertheless operate through it. Fugitive and opportunistic, the death drive seizes upon sexual desire without properly or explicitly making itself known. A sexual relation that begins with the desire to join together is interrupted by myriad forms of self-destruction that seem manifestly counter to the stated aims of the lovers. The disconcerting quality of acts that are clearly self-destructive or that destroy the bonds that one wants most to keep is but one ordinary form of wreckage by which the death drive makes itself known in sexual life.

In *Civilization and its Discontents* (1930), another text concerned with the potentially destructive effects of war, sadism is once again introduced as the 'representative' of the death drive, but in this late work the death drive is more explicitly linked with the concepts of aggression and destructiveness. These latter are no longer understood as operating exclusively in the context of sexual sadomasochism for, as Freud remarks, 'we can no longer overlook the ubiquity of non-erotic aggressivity and destructiveness' (1930: 120). Freud is registering the escalation of bellicosity and nationalism across Europe as well as the strengthening of anti-Semitism. These forms of aggression are not linked with pleasure or with the satisfactions that belong to pleasure: 'This aggressive instinct [drive] is the derivative and the main representative of the death instinct [drive] which we have found alongside of Eros and which shares world-dominion with it' (Freud 1930: 122). Even though what he now calls 'Eros' and 'Thanatos' do not usually occur separately, they nevertheless have contrary aims: Eros seeks to combine or synthesize separate units within society, bringing individuals together into groups, but also bringing groups together in the service of larger social and political forms. Thanatos drives those same units apart, and even drives each unit apart from itself. So in the very action that seeks to establish and build a social bond a counter-tendency exists that just as readily seeks to take it apart, i.e. I love you; I hate you; I cannot live without you; I will die if I continue to live with you.

Freud has two different ways of approaching this problem in relation to love. On the one hand, Freud insists throughout his work on the constitutive ambivalence in all love relations. This becomes clear in his chapter on 'emotional ambivalence' in *Totem and Taboo* (1913), but also in 'Mourning and melancholia', where the loss of the loved one is coupled with aggression (Freud 1917: 248–52). On that model, love is itself ambivalent (Freud 1917: 250). On the other hand, 'love' is another name for 'Eros' and names only one pole in the polarity of emotional ambivalence. There is love and there is hate. So either love names the ambivalent constellation of love and hate, or it is but one pole of that bipolar structure. Freud's own position seems itself to be ambivalent, perhaps rhetorically yielding further proof of his claim. Indeed, the paradoxical formulation is never fully resolved in his writings, remaining fecund throughout. It surfaces symptomatically in the late work: love is that which binds one person to another, but love, by virtue of its inherent ambivalence, contains the potential to destroy social bonds. Or, at least, if it is not love that destroys those bonds, there is a destructive force that is in love or attaches itself to love, that moves human creatures towards destruction and self-destruction, including the destruction of that which they most love.[6]

The fact that Freud's mind remains unsettled on the question of whether love contains or opposes this destructiveness is a sign of a problem that continues as he attempts to think not only about intimate relations of love but the psychology of the mass and its destructive potential. Is the destructive capacity to be found within the bonds that hold such groups together—a sort of destructive tie, or is it rather a power that 'cuts all the common bonds' (Freud 1915b: 279)—an antisocial impetus that tears at social relations?

What happens in groups, large crowds, or what Freud calls 'masses', such that destructive aims can take hold as they do? Although Freud does not lead with this question in *Group Psychology and the Analysis of the Ego* (1921), it surfaces insistently throughout the course of that text, more aptly translated as *Mass Psychology and the Analysis of the Ego*.[7] Freud notes in that text the tendency for individuals to yield their judgement or 'critical faculty' in the midst of groups, especially those compelled by a leader, and engaged in forms of mutual excitation by the prospect of violence.

The Critical Faculty and the Mass

In his discussion of Gustave Le Bon's *The Crowd: A Study of the Popular Mind*, Freud seeks to take account of what Le Bon calls 'the invincible power which allows [the individual] to yield to instincts which, had he been alone, he would perforce have kept under restraint' (Le Bon 2002: 15). The individual is 'less disposed to check himself' in an anonymous crowd (2002: 15). Freud is sceptical about whether individuals suddenly acquire new attributes in the midst of a crowd or whether, as seems more likely, what emerges are unconscious impulses that are suddenly less restrained by conscience or a sense of responsibility in the context of the group. In 1921 at least, his view is that conscience is a matter of 'social anxiety'—wondering what others think and feeling constrained by their judgements (Freud 1921: 75). A crowd that encourages its members not to worry about what others think will be relieved of inhibition and social anxiety alike.

Freud finds two interconnected descriptive terms in Le Bon quite useful for understanding what crowds do under conditions of disinhibition. The first is contagion, where the suggestion is made that contagious affect within crowds belong to an 'hypnotic order' in which the individual sacrifices his own interests for those of the collective (Freud 1921: 75). The second point is related, and it has to do with suggestibility, understood as the subjective condition of yielding to an external impression. Whereas Le Bon relies on terms like 'magnetic attraction' to explain this phenomenon, Freud remarks that those who are under the influence of suggestion—and so still within an extended region of hypnosis—undertake actions with what Le Bon calls an 'irresistible impetuosity' that is increased in the midst of crowds (Freud 1921: 76). So contagion is the central phenomenon, and suggestibility takes a more heightened form when coupled with contagion. But perhaps most importantly, the crowd is governed by unconscious motives, and those are brought forth by grand gestures: powerful images, exaggerations, and repetitions—'saying the same thing again and again' (Freud 1921: 78). The crowd expects a tendency towards violence. In Freud's words, 'it wants to be ruled and oppressed'. Precisely because Le Bon does not acknowledge any repressed content breaking out in violence, it becomes Freud's task to rethink how and why disinhibition is related to the unconscious. Freud disputes the notion that there might be hereditary grounds for the desire for an authority to which a crowd might submit. In this way, herds of animals and groups of people behave the same way: 'a group is an obedient herd' (1921: 81). Yet Freud makes clear that the mass can find a leader only if the leader himself is 'held in fascination by a strong faith (in an idea)'—only by exhibiting and communicating his own powerful belief can the leader compel the powerful belief of the mass (1921: 81). Freud writes of the leader, 'he must possess a strong and imposing will, which, the group, which has no will of its own, can accept from him' (1921: 81). In effect, his will becomes their ether—the pure essence that the gods breathe. What Freud does not immediately say, but adds later, is that as much as the mass accepts his will, the leader's will is strengthened by their belief.[8] The result of attributing such powers to a leader is the effect of 'prestige' and what Freud describes as the paralysis of 'the critical faculty' (1921: 81). Later, he remarks that within a crowd, perceiving someone else with heightened affect evokes a

sympathetic response—a form of primitive sympathy—that induces in the perceived person the same affect. This compulsion becomes stronger the more heightened the affect, and the affect heightens precisely through the sorts of contagious communications that take place in crowds. Relying on the work of William McDougall on the group mind (2015), Freud remarks that 'the most remarkable and also the most important result of the formation of a group is the "exaltation or intensification of emotion produced in every member of it"'—the result of a 'direct induction' of heightened affect (1921: 84). As a consequence, 'the individual loses his power of criticism' (Freud 1921: 84). Then again, Freud reiterates McDougall: 'the collective lowering of intellectual ability is avoided by withdrawing the performance of intellectual tasks from the group and reserving them for individual members of it'. And then Freud refines the formulation: qualities are conferred upon the mass 'which were characteristic of the individual', and those qualities are the very ones that have been effaced by the formation of the mass.

So Freud tells us that to some degree the formation of a mass makes us stupid or, rather, people congregating lose their critical faculty in concert because they have relegated the task of thought to someone else, or perhaps to no one, letting the mass replace the individual, savouring self-loss over the travails of individuation (1921: 78). Further, the heightening of affect correlates with the 'collective inhibition of intellectual functioning' (Freud 1921: 82). So though we think of mass formation as providing the condition for letting inhibitions go, the mass actually strengthens one sort of inhibition, the inhibition of critical thought. Where the inhibition of antisocial impulse once reigned, the inhibition of thought now prevails, loosening the inhibition on antisocial impulse. We have not, it seems, overcome inhibition—as seemed to be the promise—but only shifted its locus and, as it takes form as disinhibition, its operation is inverted.

Freud's view is that critical thought has the power to inhibit forms of affective excitation and inversely, heightened affect can inhibit the critical faculty. So the task that emerges for Freud is less to strengthen the powers of inhibition in general than to strengthen the inhibiting power of the critical faculty. When that heightened affect becomes a form of group cruelty, is the obligatory task to strengthen the social bonds apparently supported by love (strengthening Eros against its adversary Thanatos) or to strengthen and sharpen the critical agency that will in *The Ego and the Id* (1923) come to be identified with the superego? Here Freud is interested not only in diagnosing forms of crowd conduct that portend destruction, but also finding and strengthening whatever capacities humans have to counter such formations. Whereas love is sometimes identified as the counter-force to destruction, other times it seems it is this 'critical faculty' that is most important. In *Group Psychology* the 'critical faculty' describes various forms of deliberation and reflection, but the next year it takes on a new meaning and is associated with the superego, a form of cruelty unleashed upon the ego. Eventually, the superego will become identified as 'a pure culture of the death drive', at which point the way to counter destruction is through deliberate forms of self-restraint, i.e. directing destructiveness against one's own destructive impulse (i.e. self-restraint is a deliberate and reflexive form of destructiveness directed against the externalization of destructive aims) (Freud 1923: 53).[9] In other words, the check against unleashing destructive impulses which earlier

could be described as an inhibition is set up as a psychic mechanism once Freud introduces the superego. The task of the superego is to direct its destructive power against its destructive impulses. The problem with this solution, of course, is that an unbridled operation of the superego can lead to suicide, converting the destruction of the other into the destruction of the self. On the one hand, the 'critical faculty' seems attentive to the consequences of action, monitoring forms of expression and action to prevent injurious consequences. On the other hand, as an expression of the death drive, its aim is potentially destructive of the ego itself. A moderate form of self-checking can explode into unrestrained suicidal self-beratement, but only if the death drive itself remains unchecked. Paradoxically, this means that the critical agency upon which one relies to check destructive impulses can become the instrument of a destructive impulse, imperilling the life of the ego itself. Thus, the self-preserving tendencies of Eros have to be applied to the death drive as a check on its destructive operation. If the superego works destruction against the ego to inhibit its destructive expression, it still traffics in destruction, but the imperilled object is no longer the other or the world, but the ego itself. Thus, the critical faculty is of limited use in checking destruction, since it cannot check the destruction that operates through its superegoic form. For that, a countervailing force is needed, one that pursues self-preservation and, more generally, the preservation of life. Is that force to be called love, or is it mania? Does it involve disidentification, or the adoption of a neurotic position that establishes a critical distance from the sadistic exhilarations that run through society?

Of course, the task of checking uninhibited forms of destructiveness is not an easy one, mired as it is in certain constitutive paradoxes. If the superego were sufficient, we would have no problem, but once the superego is understood as potentially destructive as well, there seems to be no way out of the circle of destructiveness. What difference does it make if we return to the mass psychology to pose this question? For the problem is not how the superego is related to impulse (id) or the ego (reality principle), but what happens within groups: what are the perils and the promises? How does a return to this focus resituate our discussion of destructiveness?

Freud probes these questions in earnest in *Group Psychology and the Analysis of the Ego* a year or so before the development of the theory of the superego: What is the mechanism by which the disinhibition of cruelty takes place? How do we account for its workings? And what, for instance, does disinhibition have to do with identification, the process by which group identities are formed? When we say that a wave of feeling passes through a crowd, what do we mean? Or when certain kinds of passions are *unleashed* in a crowd that would otherwise remain unexpressed, how do we account for that expression? Does 'unleashing' mean that a desire was always there, but simply held in check? Or is 'unleashing' always structured, and so giving form to the desire or rage as it emerges? If we say that an elected official has licenced a new wave of misogyny, or that he has made widespread racism permissible, what sort of agency do we attribute to him? Was it there all along, or has he brought it into being? Or was it there in certain forms and now his speech and action give it new forms? In either case, impulse is structured either by the power by which it is repressed that designates and shapes it in some way or by that

power by which it is 'liberated' which endows it with specific meaning in relation to the prior repression. If we were simply to accept a hydraulic model or one that holds that a quantity of 'energy' is released when inhibition is lifted, then the impulse is the same whether it is inhibited or expressed. But if it matters through what means the inhibition has been enforced, and that means crafts the content of the repressed, then the emergence of the formerly inhibited impulse does not simply push aside the inhibiting force; rather, it wages an orchestrated attack on that form of power, debunking its reasons, its legitimacy, it claims. The impulse that emerges is thus worked over with interpretations and no raw energy, defined in relation to prohibition and licence. It has actively contested the moral and political claims that have informed and supported the inhibition; it has worked assiduously against the critical faculty—not just moral judgements and political evaluations, but the general character of reflective thought that make both possible. The impulse seeks to disperse and nullify moral self-restrictions, the basis for what Freud comes to call the superego itself. It may seem that against such a challenge to the superego, the task is to shore up moral restrictions, especially those that the self imposes on itself. But once it becomes clear that the superego is itself a potential force of destruction, the matter becomes more complex.

Freud's reflections can tell us something about what a common or shared use of the critical faculty looks like, or what it entails, what its fault lines are, and why groups and political formations can persist only when a moderating version of the critical faculty stays alive, one that can check both the external expression of destructive aims and its self-destructive internalization. The point is not to model society on the ego, but to ask about the specific social forms that destruction can take, and what implications they have for shared political life. On the one hand, Freud gives us a psychological explanation of the conditions under which groups can survive without destroying others or themselves. On the other hand, a political principle is developed that shows the importance of mass psychology for political philosophy, namely, that for societies that wish to organize themselves on the basis of a principled opposition to cruelty, that is, to war or to the death penalty, there first has to be a mindfulness of the destructive potential that inheres in all group formations, and then a shared sense of the form and power of a critical faculty, a form of judgement, that knows how and when to check that destructiveness and that has the power to do so. Such judgement is crucial to the operation of politics, and one task of political theory is to establish the conditions of its possible exercise (Zerilli 2016).

Identification, Inhibition, and the Superego

On several occasions, Freud seeks to explain that '*esprit de corps*' emerges both from envy and also from shared mourning (1921: 120). Members of a group seek to have their relations with one another mediated by an idealized leader, one they alternately admire and hate. They do not like being subordinated to the ideal they require. And even if the

idealized figure is removed, none of them can embody that ideal themselves—but they do settle on a form of equality in the wake of that loss. This is the lesson that Freud gleans from the myth of the primal horde, according to which a band of brothers murder the father only then to establish his symbolic status as a totem.[10] They bond together in a social solidarity marked by shared failure: 'No one must want to put himself forward, everyone must be the same and have the same. Social justice means that we deny ourselves many things so that others may have to do without them as well, or, what is the same thing, may not be able to ask for them. This demand for equality is the root of social conscience and the sense of duty' (Freud 1921: 121). In fact, what Freud calls 'the social feeling' depends upon 'the reversal of what was first a hostile feeling into a positively toned tie in the nature of an identification' (1921: 121). Earlier in this same text, he wrote, 'a democratic strain runs through the church—for the very reason that before Christ everyone is equal, and … everyone has an equal share of his love' (Freud 1921: 94). This potential of equality in a bond of lateral affection develops into an attachment to a shared leader or ideal of some kind, a third term that mediates potential hostility. As much as identification can occasion the suppression of hostility (and become a substitute for a libidinal attachment), it can also suppress both neurosis and the higher operations of intellectual thought. It is, of course, the latter that are required for the exercise of critical judgement—perhaps some measure of neurosis as well to the extent that it marks a withdrawal from social consensus.

Identification that depends upon an ideal—a leader, an ideal of the nation, an 'ego-ideal' of some kind—can provisionally resolve a murderous hostility, but it can also lead to the externalization of the enemy and a form of group affection that under certain conditions becomes indistinguishable from nationalism. The ideal is apparently not fully bearable—it is always superior, and envy interrupts the equality of the subjects subordinated to the one who embodies the ideal—which leads to violence against it. Indeed, when the ideal is exposed as faulty, or when it vanishes as a result of murder, the group falls asunder, and hostility also re-emerges—as does the need to establish a new ideal.

The hostility can take the form of faulting another or oneself. The excessive force of self-reprimand came within a year to be called by Freud the 'superego', at which point it became necessary to consider how best to check its operation. The 'duty' that he understood in *Mass Psychology* to be the result of a shared affection, a self-restraint imposed by individuals on themselves in the name of a greater good, is redescribed in *The Ego and the Id* as the superego (Freud 1923). And, as already noted in the 'The critical faculty and the mass', it is at this point that the critical faculty takes on a new and dangerous character. It is no longer simply the faculty of a reason that checks impulse, but the instrument of a destructive power and unbridled sadism. The superego has within it a 'destructive component' that turns against the ego, finds it faulty, even threatens it with lethal punishment (Freud 1923: 43). The 'critical agency' now identified with the superego imposes restraint and restriction just like a strong inhibition would do, except that it also runs the risk of a fuller destruction of the ego. Freud puts the matter this way:

… the excessively strong super-ego which has obtained a hold on consciousness rages against the ego with merciless violence, as if it had taken possession of the whole of the sadism available in the person … What is now holding sway in the super-ego is, as it were, a pure culture of the death drive, and in fact it often enough succeeds in driving the ego into death (1923: 53).

What, if anything, checks the merciless violence of one part of the self unleashed against another? Freud finishes that sentence by claiming that what can thwart the success of self-destruction is for the ego to 'fend off its tyrant in time by the change round into mania' (1923: 53).

Freud here references his 1917 work on 'Mourning and melancholia' where he seeks to distinguish between mourning, which implies a wakeful acknowledgement of the reality of a lost person or ideal, and melancholia, which is a failure to acknowledge the reality of loss. In melancholia the lost other is internalized (in the sense of incorporated) as a feature of the ego, and a form of heightened self-beratement re-enacts—and inverts—at a psychic level the relation of the ego to the lost other. The recrimination against the lost person or ideal is 'turned round' against the ego itself, and in this way the relation is preserved as an animated intrapsychic relation (Freud 1917: 251). Even in that essay, Freud makes clear that the hostility unleashed against the ego is potentially fatal. The scene of melancholic self-beratement thus becomes the model for the later topography of superego and ego. Melancholia, however, is composed of two opposing trends: the first is self-beratement which becomes the signature action of 'conscience'; the second is 'mania' which seeks to break the bond to the lost object, actively renouncing the object that is gone (Freud 1917: 253–5). The 'manic' and energetic denunciations of the object, the ego's heightened efforts to break the bond to the lost object or ideal, imply the desire to survive the loss and not to have one's own life claimed by the loss itself. Mania is, as it were, the protest of the living organism against the prospect of its destruction by an unchecked superego.

As we have seen, forms of heightened self-beratement (and, by extension, self-destruction) can be provisionally ameliorated through the social solidarity of failure in which none of us lives up to the ideal, and this shared failure grounds our solidarity and sense of equality. That amelioration of superegoic violence proves to be provisional when a group formation fails to organize and contain that hostility, and it can assume a fatal form. Moreover, there are group formations that mobilize that destructive hostility towards an external enemy, at which point the destruction of life, even the mass destruction of life, becomes possible. Significantly, then, when Freud identifies 'mania' as the force by which the internal tyrant might be overthrown, he is referring to that power of the organism to break the bonds of identification itself. Indeed, active *disidentification* becomes one way of the organism to ameliorate its self-destructive powers in order precisely to survive. Mania serves the purposes of disidentification, but it also presumes a 'criticism' of the tyrant, offering one way to think about the exercise of a *critical faculty that is directed against the superego* and so not already identified in its terms.

In brief, the superego relied upon to check destructive impulse is always potentially an instrument of destruction, which is not only why violence can be moralized, deemed as righteous, but also why those with particularly severe consciences can take their own life. Similarly, it may seem that identifications are important for empathy and the perpetuation of social bonds, but they also imply destructive potential. The internalization of the lost other or ideal is one form of identification, one that preserves and animates a form of hostility that has the power to destroy the living organism itself. So, even as the superego checks the externalization of destructiveness, it remains a potentially destructive instrument that can come to serve the very murderous purposes it is meant to check—but mainly through self-murder. We return again to the question: Is there a way to oppose destructiveness that does not escalate destructiveness? It would seem that the superego is a weak instrument to enforce a check on violence, unless we opt for the violence of the superego, however fatal that may prove, over its alternative, externalized expression. But is that the only path available?

Freud and Einstein: Solidarity in Life

This very problem emerges in the correspondence between Freud and Einstein in 1931–32, the years directly preceding Hitler's rise to power and their subsequent exile from Germany and Austria respectively.[11] Einstein writes to Freud to ask him how humankind can become delivered from 'the menace of war' (Freud 1933b: 199). Lamenting that the fate of humankind is in the hands of 'a governing class' that is 'craving for power', and 'hostile to any limitation of national sovereignty', Einstein appeals to Freud as someone whose 'critical judgement' is most important at this time in which world war once again threatens Europe. He asks whether there is a basis in the drives that constitute human psychic life for a political arrangement that could serve as an effective check on war. In particular, he asks whether it might be possible to establish an association or a tribunal that could check the destructive power of those drives. Einstein first identifies the problem as destructive drives, but he is also interrogating the issue at the level of political institutions, calling for nations to cede their sovereignty to an international body that would demand a commitment to preventing war and guaranteeing international security. This political goal can only be achieved if human beings are the kind of creatures capable of constituting, and submitting to, international authorities that have the power to prevent war. If there is a tendency or drive that undercuts that capacity, then averting war may well be impossible. Clearly having read Freud, Einstein asks whether human beings have within them 'a lust for hatred and destruction' and whether this can be 'raise(d) … to the power of a collective psychosis' (Freud 1933b: 201). So though he wonders whether the destructive drives can be contained, he is also wondering whether human practices of institutions can be cultivated that would increase the possibility of preventing war. He notes that violence can take the form of wars between nations but

also civil wars motivated by religious zealotry, as well as 'the persecution of racial minorities' (Freud 1933b: 201).

Freud's response is complex, beginning with a report on his initial response to Einstein's letter: he is not a 'statesman' and daunted by practical matters (1933b: 203). He then retracts his initial claim not because he believes he *is* a political thinker but because, with Einstein, he shares a political desire for peace and for the alleviation of suffering, and so must be engaged in political reflections as a 'psychological observer'. Freud warns that he has no practical proposals, but his remarks do elaborate a political position with both theoretical and practical implications. His first proposal is to replace a distinction Einstein makes between right and power with one between right and violence ('right' translates '*Recht*' which, in German, means 'legal order' and even 'justice'). In Freud's account, conflicts between persons and groups have been traditionally resolved through recourse to violence, but this happens less regularly as group formations change. In particular, he notes that 'a path was traced that led away from violence to law' when 'an alliance of many weaklings' overcomes the strength of the single man or leader. In this way, he writes, 'brute force is overcome through union' or what he also calls 'the power of a community'. In his view, 'the superior strength of a single individual could be rivalled by the union of several weak ones' and later he elaborates: '… in order that the transition from violence to this new right or justice may be effected … [t]he union of the majority must be a stable and lasting one'. To do this, a psychological condition has to be met: 'the growth of … communal feelings which are the true source of its strength'.

Writing to Einstein a full decade after *Mass Psychology*, Freud now conjectures that the community is held together *not* by their common subordination to an ideal leader, but precisely through their explicit power to overthrow a tyrant or authoritarian ruler and to establish common and enforceable laws and institutions in the wake of that overthrow. Can the 'communal feelings' and 'emotional ties' actually persist to achieve such a goal? To answer this political question, the turn to psychological conditions once again becomes obligatory. The answer seems to depend on how we interpret the 'community of interests' (Freud 1933b: 205). Freud's wager is that as power (not violence) is transferred to ever larger combinations, group members are increasingly enfranchised and more inclined to act from sentiments of solidarity. Einstein talked about the obligation of each nation state to surrender its sovereignty to a larger international body. Freud also imagines the distribution of power beyond the model of sovereignty. As the community and its powers of self-governance expand and become increasingly distinct from, even opposed to, the individual ruler, the sentiment of solidarity, expressed in a set of laws both self-legislated and self-restraining, is relied upon to check destructiveness. The ongoing problem, however, is that violence can erupt within the community when, for instance, one faction pits itself against another or the right of rebellion is exercised against the state or the international body that limits the sovereignty of states.

The limitation on violence seems to coincide, for both Freud and Einstein, with the limitation of state sovereignty within a broader internationalist frame. In the 1930s, both Freud and Einstein understood nationalist fervour to lead to outbreaks of violence,

though neither could fully see the forms of state violence in fascism and Nazism that would materialize in the next few years. The international body or 'tribunal' they both imagined was to some degree represented by the League of Nations in the early 1930s, but that institution hardly constituted an ultimate power since state sovereignty was not effectively checked by membership. Ideally, an international institution for developing and maintaining international law with the power to enforce its judgements would provide a check on war, but only, of course, if international sentiments of solidarity could be maintained to support its existence and function. Without the power of enforcement, such a body lacks the sovereign power it requires to prevent war. The internationalist wager that both Freud and Einstein considered was that international institutions could be based on modes of self-governance across national borders that delocalize and limit sovereign power. Einstein, who called himself 'immune from nationalist bias', thought that the risk was worth taking: 'the quest for international security involves the unconditional surrender by every nation, in a certain measure, of its liberty of action, its sovereignty that is to say, and it is clear beyond all doubt that no other road can lead to such security'. He then continues to remark upon the failure of this effort, which 'leaves us no room for doubt that strong psychological factors are at work' (Freud 1933b: 200).

Freud is, of course, aware that law can work in several ways, and that it makes no sense to have a naïve faith in the power of law. It can encode and reproduce social inequalities, and more often than not 'laws are made by and for the rulers'. The only way for law not to become an instrument that fortifies the power of the rulers over the people is for the masses to undergo what he calls 'a cultural evolution' such that the ideals embodied in international self-governance become increasingly desirable and widespread (Freud 1933b: 214). If the ideals of the community become nationalistic, then violence returns in the form of war among competing nation states. Thus, internationalist ideals support the sentiment of solidarity for Freud. If, however, the idea of internationalism still imagined communities and groups primarily as nations, it may be that 'transnational' solidarity proves to be the better name for sentiments of solidarity no longer structured by the nation state and its corresponding nationalism. But such a notion postdates Freud.

The problem, however, remains whether such sentiments can fully and effectively contain or redirect aggression. Freud has offered some paths for providing an answer, but he does not follow through on all of the possibilities that his work has opened up for consideration. In conclusion, I would like to consider Freud's final case for ways of preventing war as it takes him on a train of thought unpursued in his reflections on group psychology; the first requires resisting the exhilarations of nationalism; the second, heeding the *organic* basis of our nature as human beings. Although the drive is not the same as the instinct, it does not follow that the drive is relieved of all organic dimensions. Indeed, persistence always takes some form or another, and yet it is not reducible to any of its possible forms. That paradox informs one important argument Freud makes for an internationalist resistance to war.

As we have seen, for Freud, there is no possibility of finally 'conquering' the death drive and its forms of aggression and destructiveness. Towards the end of his letter to Einstein, he makes a strong case that there are only two ways to counter the propensity for war: the mobilization of its 'Eros, its antagonist', and the forming of communal forms of identification (Freud 1933b: 212). A third possibility was suggested by Einstein, namely, that an elite class of intellectuals should form an association that could provide proper guidance for the masses. Freud does not readily accept this suggestion, maintaining instead that an evolution of the masses is possible through education and the cultivation of solidaristic sentiments of a non-nationalist sort.[12] The ideal condition would be one in which every member of a community exercises self-restraint, and does so precisely by recognizing that the preservation of life is itself a good to be valued in common. Freud realizes that this is utopian, but it provides a subtle counterpoint to Einstein's more frank embrace of elitism.

Although Freud does not pursue the political alternative to elitism, his ideal of a community whose members are equally bound to impose self-restraint in the name of the preservation of life opens the possibility of the democratization of critical judgement and critical thought that does not rush unimpeded to the extreme of superegoic self-flagellation. Does he in the end offer a strong enough response to the sceptical position that the destructive powers of humans are so profoundly inscribed in the life of the drives that no political arrangement can effectively check them? On the one hand, Freud argues that we must ally behind love, which builds and preserves social bonds, and identification, which builds and preserves sentiments of solidarity, over and against hate, or Thanatos, which tears at social bonds in wild and mindless ways. On the other hand, he has time and again underscored the fact that love and hate are equally constitutive dimensions of the drives, and that it is not possible to eliminate destructiveness through amplifying Eros. It is not only that we must sometimes *aggressively* defend our lives in order to preserve life (the aim of Eros); we also have to commit to living with those towards whom we maintain intense feelings of hostility. In his discussion of identification and melancholia, it is clear that all love relations contain ambivalence, which suggests that love and hate are both equally constitutive of love. So love names one pole in the oppositional relation of love and hate. But it also names the opposition itself, lived out as emotional ambivalence. One can say that I love you and so do not hate you, but one can also say that love and hate are bound together, and this paradox is what we name by 'love'. In the former formulation, love is unequivocal; in the latter formulation, love does not escape ambivalence. Is there something about the rhythm, however jarring, established between these two formulations that constitutes a broader concept of love for Freud?

There seem to be two consequences then to Freud's views on destructiveness and war that that are opened up but not precisely pursued; both pertain directly to the political philosophy embedded in his reflections on group (or mass) psychology. The first is that a corrective to forms of accelerated nationalist sentiment is precisely ambivalence, the 'tearing' at the social bond that follows from a mindful self-distancing from its exhilarations and hostilities—and from the restrictively nationalist framework. One might

Subscriber: Tulane University; date: 15 January 2019

at the same time love a country and dissent from its nationalist fervour—that would activate ambivalence in the service of a critical reflection on the possibility of war and a refusal to partake in its excitations. The second would be to rally hatred against war itself. Freud offers this indirectly in his letter to Einstein in his own rhetoric. For instance, he writes, 'the basis for our common hatred of war … is that we cannot do otherwise than to hate it. Pacifists we are because our organic nature wills us thus to be' (Freud 1933b: 214). Significantly, it is as a result of the 'growth of culture' or what he earlier called its 'evolution' that we come to resent war and find it intolerable. War sensations, he claims, no longer thrill us, only because we have come to see—and to imagine—the destruction of organic life that war implies, something which is unbearable for humans to accept in light of their own organic life. On the one hand, it is organic life that makes us pacifists, since we do not will our own destruction (when we are not under the sway of the death drive). On the other hand, we only come to understand the consequences of the destruction of organic life through a cultural process that allows us to see and consider this destruction and so to develop a revulsion against destruction itself. In the end, Freud hopes that organic life will have the final say against the death drive whose aim is the destruction of that very life, and that various forms of organic life will be understood as connected through relations of dependency that extend throughout the living world. His is a politics of and for the living organism, even if sometimes the organism is swayed by the circuitous or destructive path. Hatred is never fully absent, but its negative power can become focused as an aggressive stance against war, one form of destruction pitted against another, a view that would be compatible, for instance, with an aggressive form of pacifism, what Einstein himself called 'militant pacifism'.[13]

So if the superego was established by Freud as that internal mechanism fuelled by the death drive, constituting a pure culture of the death drive, and its limit point was the destruction of the ego and the living organism itself (suicide or murder), then the form of aggression that Freud imagines at the end of his correspondence with Einstein is of a different order. When he remarked that the only hope for prevailing against the murderous aim of the unbridled superego which would, as it were, judge the ego unto death, was the 'manic' possibility of breaking with the tyrant (levelling plaint after plaint), he offered us a glimpse into those forms of insurrectionary solidarity that turn against authoritarian and tyrannical rule, exceed the national frame, and seek binding international agreement. The hatred directed against war is perhaps like the mania that alone has the strength to free the subject from the tyrant; both break with nationalist and militarist forms of social belonging through turning one sense of the critical faculty against another. The critical faculty that becomes animated in the name of a democratization of dissent is one that opposes war and resists the intoxications of nationalism, turns against the leader who insists that obedience to a warmongering authority is obligatory. In this way Freud offers us the possibility of imagining the democratization of critical judgement based on sentiments of solidarity, one that turns against that life-threatening form of aggression, including its critical manifestation. Aggression and hatred both remain, for sure, but directed now against all that which undermines the prospect of expanding equality and imperils the organic persistence of

our interconnected lives. But nothing remains guaranteed, for the death drive appears as well to be part of organic life, so if the organic turns out to be driven by the duality of life and death, that should hardly come as a total surprise. The struggle that constitutes us as political creatures is the one we continue in the practices of life and death without a perfect conscious understanding despite our occasionally admirable and decisive efforts at vigilance.

References

Alford, C. F. (1994). *Group Psychology and Political Theory*. New Haven, CT: Yale University Press.

Balibar, É. (2017). 'The invention of the superego: Freud and Kelsen, 1922'. In *Citizen Subject: Foundations for Philosophical Anthropology*, trans. S. Miller (pp. 227–255). New York: Fordham University Press.

Bottici, C. and Kuehner, A. (2012). 'Between psychoanalysis and political philosophy: Towards a critical theory of political myth'. *Critical Horizons 13*(1): 94–112.

Brown, W. (2001). 'The desire to be punished: Freud's "A child is being beaten"'. In *Politics Out of History* (pp. 45–61). Princeton, NJ: Princeton University Press.

Freud, S. (1913). *Totem and Taboo*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. vii–162). London: Hogarth Press.

Freud, S. (1915a). 'Instincts and their vicissitudes'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 109–140). London: Hogarth Press.

Freud, S. (1915b). 'Thoughts for the times on war and death'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 273–300). London: Hogarth Press.

Freud, S. (1917). 'Mourning and melancholia'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 14 (pp. 237–258). London: Hogarth Press.

Freud, S. (1920). *Beyond the Pleasure Principle*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 18 (pp. 1–64). London: Hogarth Press.

Freud, S. (1921). *Group Psychology and the Analysis of the Ego*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 18 (pp. 65–144). London: Hogarth Press.

Freud, S. (1923). *The Ego and the Id*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 19 (pp. 1–59). London: Hogarth Press.

Freud, S. (1924). 'The economic problem of masochism'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 19 (pp. 155–170). London: Hogarth Press.

Freud, S. (1930). *Civilization and its Discontents*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 21 (pp. 57–146). London: Hogarth Press.

Freud, S. (1933a). *New Introductory Lectures on Psycho-Analysis*. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 22 (pp. 1–182). London: Hogarth Press.

Freud, S. (1933b). 'Why war?'. In J. Strachey (ed. and trans.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 22 (pp. 195–216). London: Hogarth Press.

Freud, S. (2004). *Mass Psychology and the Analysis of the 'I'*. In J. A. Underwood (trans.), *Mass Psychology and Other Writings*. New York: Penguin. (Original work published in 1921.)

Frosh, S. (1999). *The Politics of Psychoanalysis*: *An Introduction to Freudian and Post-Freudian Theory*. London: Palgrave.

Gerson, G. (2004). 'Object relations psychoanalysis as political theory'. *Political Psychology 25*(5): 769–794.

Holton, G. J. and Elkana, Y. (eds) (1982). *Albert Einstein: Historical and Cultural Perspectives*. New York: Dover.

Kafka, F. (1971). 'Josephine the Singer, or the Mouse Folk'. In N. Glatzer (ed.), *Complete Stories*, trans. W. Muir and E. Muir (pp. 386–403). New York: Schocken Books. (Original work published in 1924.)

Lear, J. (1998) *Open-Minded: Working Out the Logic of the Soul*. Cambridge, MA: Harvard University Press.

Le Bon, G. (2002). *The Crowd: A Study of the Popular Mind*. New York: Dover. (Original work published in 1895.)

Luxon, N. (2013). *Crisis of Authority: Politics, Trust, and Truth-Telling in Freud and Foucault*. Cambridge: Cambridge University Press.

McDougall, W. (2015). *The Group Mind*. Cambridge: Cambridge University Press. (Original work published in 1920.)

Pitkin, H. F. (1999). *Fortune Is a Woman: Gender and Politics in the Thought of Niccolo Machiavelli*. Chicago, IL: University of Chicago Press.

Reich, W. (1970). *The Mass Psychology of Fascism*. New York: Farrar, Strauss, Giroux. (Original work published in 1933.)

Roazen, P. (1999). *Freud: Political and Social Thought*. New Jersey: Transaction Publishers.

Roazen, P. (2000). *Political Theory and the Psychology of the Unconscious: Freud, J. S. Mill, Nietzsche, Dostoevsky, Fromm, Bettelheim and Erikson*. London: Open Gate Press.

Rogin, M. (1975). *Fathers and Children: Andrew Jackson and the Subjugation of the American Indian*. New York: Knopf.

Rogin, M. (1988). *Ronald Reagan: The. Movie, and Other Episodes in Political Demonology*. Berkeley, CA: University of California Press.

Rose, J. (1993). *Why War? Psychoanalysis, Politics, and the Return to Melanie Klein*, London: Blackwell.

Rose, J. (2004). 'Introduction'. In S. Freud, *Mass Psychology and Other Writings* (pp. vii–xlii). New York: Penguin.

Rose, J. (2007). *The Last Resistance*. London: Verso.

Rosen, A. (2007). 'On the fate of psychoanalysis and political theory'. *Psychoanalytic Quarterly 76*(3): 943–980.

Rubenstein, D. (2008). *This is Not a President: Sense, Nonsense, and the American Political Imaginary*. New York: New York University Press.

Wolfenstein, E. V. (1993). *Psychoanalytic-Marxism: Groundwork*. New York: Guilford Press.

Wolin, S. (2004). 'Liberalism and the decline of political philosophy'. In *Politics and Vision: Continuity and Innovation in Western Political Thought* (pp. 257–314). Princeton, NJ: Princeton University Press. (Original work published in 1960.)

Yanay, N. (2014). *The Ideology of Hatred: The Psychic Power of Discourse*. New York: Fordham University Press.

Zaretsky, E. (2015). *Political Freud: A History*. New York: Columbia University Press.

Zerilli, L. (2016). *A Democratic Theory of Judgment*, Chicago, IL: University of Chicago Press.

Notes:

(^{1}) [Eds: For a discussion of developments in psychoanalytic thought on the death drive and its containment, see Frosh (this volume), § 'War, barbarism, rationality'.]

(^{2}) Freud writes in 'Instincts and their Vicissitudes': '… an "instinct" appears to us as a concept on the frontier between the mental and the somatic…' […so erscheint uns der 'Trieb' als ein Grenzbegriff zwischen Seelischem und Somatischem…"(1915: 121–22). Freud, *Psychologie des Unbewussten*, Fischer Verlag, 85.]

(^{3}) In *Beyond the Pleasure Principle*, Freud reassures his reader that there is no reason to worry about the speculative character of his discussion of the drives, since 'we are obliged to operate with the scientific terms, that is to say with the figurative language, proper to psychology …' (1920: 60). In the *New Introductory Lectures*, he goes further: 'The drives are so to say our mythology. Drives are mythical entities, magnificent in their indefiniteness' (1933a: 95).

(^{4}) In *Beyond the Pleasure Principle*, Freud describes the 'circuitous route to death' that a life takes, governed in fugitive fashion by a drive that seeks to return to an inorganic condition (1920: 38).

(^{5}) Freud's theorization of sadomasochism seeks to explain the phenomenon through recourse to the theory of libido in 'Instincts and their Vicissitudes' (1915a) but is revised in light of the death drive in *Beyond the Pleasure Principle* (1920) and 'The Economic Problem of Masochism' (1924).

(^{6}) [Eds: see Backström (this volume) for a critical discussion of Freud's conception of love that argues that ambivalent forms of love are failures in love.]

(^{7}) In German, it is not group psychology, but 'mass psychology' [*Massenpsychologie*], which carries a sense of much larger crowds, drawing on Le Bon's analysis of the crowd. In the Penguin edition published in 2004 and translated by J. A. Underwood, the text is appropriately entitled *Mass Psychology and the Analysis of the 'I'*.

(^{8}) See for instance Franz Kafka's 'Josephine the Singer, or the Mouse Folk' for an account of how the leader—a singer—is inflated by the people, idealized in impossible terms, because they depend on her for a sense of their national unity and belonging. In the end, she is weak and barely speaks or sings, but the mechanism of their idealization stays intact, and it appears that the people (the mouse folk) are the ones with the idealizing power to sustain their leader. Finally, however, transience takes over, and the entire history is ushered into oblivion.

(^{9}) Translation modified.

(¹⁰) Freud considers the myth of the 'primal horde' in the postscript to *Mass Psychology*, an origin story in which the original Father is killed and dismembered by the sons who idealize him. Idealization and murderous impulse work together—a primary form of ambivalence that is linked to Freud's views on the idealizing and murderous dimensions of love. The totemic brotherhood that comes about after the kill mourns the loss, but reaps its benefits in a form of political equality among the brothers. But equality is apparently hard to bear, and the ideal of the father re-emerges now within the family, where the husband's power breaks the line of matriarchal rule. Their collective act of murder left them fatherless, but with the institution of familial patriarchy through the categories of 'husband' and 'father' they each become the substitute for the father-ideal they have lost (Freud 1921: 135; Freud 2004: 91).

(¹¹) Einstein departed Germany in 1933 and Freud left Vienna in 1938. Their correspondence can be found as 'Why War?' (Freud 1933b). In 1931, the Institute for Intellectual Cooperation invited Einstein to engage in a dialogue with a thinker of his choice on the topics of politics and peace, and he chose Freud, whom he had met briefly a few years before.

(¹²) For Freud's resistance to nationalism and Zionism, see Jacqueline Rose, *The Last Resistance* (2007: 17–38).

(¹³) See Albert Einstein's interview with George Sylvester Viereck in January, 1931 where he claims, 'I am not only a pacifist, but a militant pacifist. I am willing to fight for peace. Nothing will end war unless the peoples themselves refuse to go to war. Every great cause is first championed by an aggressive minority' (quoted in Holton and Elkana 1982: 377).

Judith Butler

Comparative Literature, University of California at Berkeley